GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:02:56 ; Search time 38 Seconds
(without alignments)
578.588 Million cell updates/sec

Title: US-09-830-964-1
Perfect score: 846
Sequence: 1 APPRLICDSRVLERYLLEAK.....SNFLRGKLKLYTGEACRTGD 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	846	100.0	165	21	AAY94605	Human erythropoiet
2	846	100.0	165	21	AAB03760	Human erythropoiet
3	846	100.0	165	21	AAY99705	Non-glycosylated e
4	846	100.0	165	21	AAY93445	Amino acid sequenc
5	846	100.0	165	22	AAB84525	Amino acid sequenc
6	846	100.0	165	22	AAB66697	Human erythropoiet
7	846	100.0	165	23	ABB77896	Amino acid sequenc
8	846	100.0	165	23	AAM53061	Human erythropoiet
9	846	100.0	166	8	AAP70398	Sequence of human
10	846	100.0	166	13	AAR23593	Recombinant hemato
11	846	100.0	166	19	AAW77780	Human EPO receptor
12	846	100.0	166	19	AAW58404	Human erythropoiet
13	846	100.0	166	21	ABB07030	Modified erythropo
14	846	100.0	166	22	AAE02641	Human erythropoiet
15	846	100.0	166	22	AAB66698	Human erythropoiet
16	846	100.0	166	23	ABB77897	Amino acid sequenc
17	846	100.0	166	23	AAM53062	Human erythropoiet
18	846	100.0	167	6	AAP50298	Human recombinant
19	846	100.0	167	6	AAP50299	Human recombinant
20	846	100.0	169	23	ABB77899	Amino acid sequenc
21	846	100.0	174	23	ABB77898	Amino acid sequenc
22	846	100.0	174	23	ABB77900	Amino acid sequenc
23	846	100.0	188	7	AAP60599	Clone lambda HEPOF
24	846	100.0	188	9	AAP81195	Erythropoietin enc
25	846	100.0	193	6	AAP50300	Human erythropoiet
26	846	100.0	193	7	AAP60597	Clone lambda HEPOF
27	846	100.0	193	8	AAP70256	Sequence of human
28	846	100.0	193	15	AAR65499	Human prepro-eryth
29	846	100.0	193	16	AAR81982	Human erythropoiet
30	846	100.0	193	16	AAR71137	Human erythropoiet
31	846	100.0	193	16	AAR74141	Human erythropoiet
32	846	100.0	193	17	AAR98397	Human erythropoiet
33	846	100.0	193	21	AAY94530	Human erythropoiet
34	846	100.0	193	21	AAY93638	Amino acid sequenc
35	846	100.0	193	21	AAY99704	Human non-glycosyl
36	846	100.0	193	21	AAY43398	Human erythropoiet
37	846	100.0	193	22	AAB85573	Human erythropoiet
38	846	100.0	193	22	AAB34978	Human erythropoiet
39	846	100.0	193	23	AAE15341	Human erythropoiet
40	846	100.0	194	16	AAR71167	Human erythropoiet
41	846	100.0	194	19	AAW62048	Human erythropoiet
42	846	100.0	194	21	AAB10654	Human erythropoiet
43	846	100.0	196	23	ABB77902	Amino acid sequenc
44	846	100.0	201	23	ABB77901	Amino acid sequenc
45	846	100.0	201	23	ABB77903	Amino acid sequenc

ALIGNMENTS

RESULT 1
AAY94605
ID AAY94605 standard; Protein; 165 AA.
XX
AC AAY94605;
XX
DT 28-NOV-2000 (first entry)
XX
DE Human erythropoietin.
XX
KW Human; erythropoietin; EPO; purification; anaemia.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 24
FT /note= "N-Glycosylation site"
FT Modified-site 38
FT /note= "N-Glycosylation site"
FT Modified-site 83
FT /note= "N-Glycosylation site"
FT Modified-site 126
FT /note= "O-Glycosylation site"
XX
PN WO200027869-A1.
XX
PD 18-MAY-2000.
XX
PF 08-NOV-1999; 99WO-US26241.
XX
PR 06-NOV-1998; 98AR-0105610.
PR 23-FEB-1999; 99AR-0100680.
XX

PA (STER-) STERRENBELD BIOTECHNOLOGIE NORTH AMERICA.
XX
PI Carcagno CM, Criscuolo M, Melo C, Vidal JA;
XX
DR WPI; 2000-376485/32.
XX
PT Novel methods for purifying recombinant human erythropoietin from
PT mammalian cell culture reagents -
XX
PS Claim 16; Page 18; 30pp; English.
XX
CC The present invention relates to a method for purifying erythropoietin
CC (EPO) for treatment of disease, especially anaemia. The method involves
CC treating cell culture supernatants with differential precipitation,
CC hydrophobic interaction chromatography, diafiltration, anionic and
CC cationic exchange chromatography and molecular exclusion
CC chromatography. The present sequence is the protein from the culture
CC supernatant of transfected cell lines, after purification by the above
CC process. The sequence shows total homology with natural human EPO.
CC The advantage of this method is that high purity and quality EPO is
CC produced. A further advantage is that the process does not involve the
CC use of organic solvents that may harm the environment.
XX
SQ Sequence 165 AA;

Query Match 100.0%; Score 846; DB 21; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
Db 1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy 61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
Db 61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy 121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
Db 121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165

RESULT 2
AAB03760
ID AAB03760 standard; protein; 165 AA.
XX
AC AAB03760;
XX
DT 04-OCT-2000 (first entry)
XX
DE Human erythropoietin (EPO) amino acid sequence.
XX
KW Erythropoietin; EPO; human; erythroblast differentiation; anaemia;
KW large scale production; renal failure.
XX
OS Homo sapiens.
XX
PN WO200027997-A1.
XX
PD 18-MAY-2000.
XX
PF 08-NOV-1999; 99WO-US26240.
XX
PR 06-NOV-1998; 98AR-0105611.
PR 23-FEB-1999; 99AR-0100681.
XX
PA (STER-) STERRENBELD BIOTECHNOLOGIE NORTH AMERICA.
XX
PI Carcagno CM, Criscuolo M, Melo C, Vidal JA;
XX
DR WPI; 2000-376519/32.
XX
PT A novel method for the massive culture of recombinant mammalian cells
PT producing recombinant human erythropoietin -
XX
PS Example 8; Page 11-12; 23pp; English.
XX
CC This sequence represents the human erythropoietin amino acid sequence.
CC Erythropoietin is a glycoprotein that stimulates erythroblast
CC differentiation in the bone marrow. The present invention relates to a
CC method for the large scale production of human EPO from recombinant
CC mammalian cells. The method comprises culturing mammalian cells which
CC express recombinant human EPO in culture medium comprising insulin.
CC Erythropoietin can be used to treat anaemia derived from renal failure.
CC The method allows for the industrial scale production of EPO, and
CC overcomes the problems of low reproducibility and output quality which
CC are encountered with previous production methods.
XX
SQ Sequence 165 AA;

Query Match 100.0%; Score 846; DB 21; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
Db 1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy 61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
Db 61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy 121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
Db 121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165

RESULT 3
AAY99705
ID AAY99705 standard; Protein; 165 AA.
XX
AC AAY99705;
XX
DT 15-SEP-2000 (first entry)
XX
DE Non-glycosylated erythroprotein analogue NGE-166delta.
XX
KW Human; non-glycosylated erythroprotein analogue; NGEA; haematocrit;
KW antianaemic; anaemia; erythropoiesis promoter; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200032772-A2.
XX
PD 08-JUN-2000.
XX
PF 23-NOV-1999; 99WO-US27801.
XX
PR 30-NOV-1998; 98US-0110289.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Beals JM, Glaesner W, Micanovic R, Millican RL, Witcher DR;
XX
DR WPI; 2000-412320/35.
DR N-PSDB; AAA48373.
XX
PT Non-glycosylated erythropoietic compound useful for increasing
PT hematocrit level in mammal with insufficient hematocrit levels in
PT conditions such as anemia, comprises protein covalently bonded to
PT polymer -
XX
PS Claim 2; Page 93-94; 94pp; English.
XX
CC The present sequence is a non-glycosylated erythroprotein analogue

```
CC   (NGEA) designated NGE-166delta. The protein sequence is identical to
CC   the sequence of wild-type human non-glycosylated erythroprotein NGE
CC   except that Arg at position 166 is deleted. NGE promotes erythropoiesis
CC   and can therefore be used to increase haematocrit levels in mammals
CC   with conditions such as anaemia, in which levels of haematocrit are
CC   insufficient. NGE analogues can also be used to treat such conditions.
CC   NGEAs do not themselves cause a significant increase in haematocrit but
CC   they acquire that property once they are derivatised with polyethylene
CC   glycol polymers. The analogues can be produced using a linkerless
CC   aldehyde modification process. They show stability and bioactivity in
CC   vivo. The nucleotide sequence encoding this protein was constructed
CC   synthetically by in vitro hybridisation using a set of six overlapping
CC   oligonucleotides from the positive strand of human erythropoietin cDNA
CC   with six complementary oligonucleotides (negative strand). The codon
CC   usage was 100% optimised for E. coli codon usage. The hybridised
CC   oligonucleotides were ligated with T4 DNA ligase and the ligation product
CC   amplified by PCR. The nucleotide sequence was used to express the protein
CC   in host cells.
XX
SQ   Sequence   165 AA;

  Query Match             100.0%;  Score 846;  DB 21;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-86;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 4
AAY93445
ID   AAY93445 standard; protein; 165 AA.
XX
AC   AAY93445;
XX
DT   04-SEP-2000  (first entry)
XX
DE   Amino acid sequence of human erythropoietin.
XX
KW   Human; erythropoietin; EPO; anaemia; renal failure.
XX
OS   Homo sapiens.
XX
PN   WO200028066-A1.
XX
PD   18-MAY-2000.
XX
PF   08-NOV-1999;   99WO-US26238.
XX
PR   06-NOV-1998;   98AR-0105609.
PR   23-FEB-1999;   99AR-0100679.
XX
PA   (STER-) STERRENBELD BIOTECHNOLOGIE NORTH AMERICA.
XX
PI   Carcagno CM,  Criscuolo M,  Melo C,  Vidal JA;
XX
DR   WPI; 2000-376574/32.
XX
PT   New host cell producing recombinant human erythropoietin (EPO) used for
PT   large scale production of EPO  -
XX
PS   Claim 1; Page 26-27; 51pp; English.
XX
CC   The present sequence represents human erythropoietin protein. The
CC   specification describes a host cell line which is used to produce
CC   human erythropoietin (EPO). EPO is a glycoprotein. The cell line
CC   is used for the production of recombinant human erythropoietin. The
CC   protein is used for the treatment of anaemia, especailly anaemia derived
CC   from renal failure.
XX
SQ   Sequence   165 AA;

  Query Match             100.0%;  Score 846;  DB 21;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-86;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 5
AAB84525
ID   AAB84525 standard; protein; 165 AA.
XX
AC   AAB84525;
XX
DT   05-SEP-2001  (first entry)
XX
DE   Amino acid sequence of human erythropoietin (EPO) protein.
XX
KW   Erythropoietin; EPO; erythropoietin stimulating protein; NESP;
KW   sustained release.
XX
OS   Homo sapiens.
XX
PN   WO200130320-A1.
XX
PD   03-MAY-2001.
XX
PF   23-OCT-2000; 2000WO-US29257.
XX
PR   22-OCT-1999;   99US-0426566.
PR   13-OCT-2000; 2000US-2222222.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Burke P,  Klumb L,  Murphy K,  Herberger J,  French DL;
XX
DR   WPI; 2001-417552/44.
XX
PT   Sustained release composition comprises an active biological
PT   ingredient, notably a protein or other biopolymer, particularly
PT   erythropoietin stimulating protein, in biocompatible, biodegradable
PT   polymeric microparticles  -
XX
PS   Disclosure; Page 56; 61pp; English.
XX
CC   The present sequence encodes a human erythropoietin (EPO) protein.
CC   The specification describes a composition for the sustained release of
CC   biologically active EPO stimulating protein (NESP). The reduced
CC   frequency of administration of NESP, which requires preferably injection
CC   by skilled personnel, improves patient compliance. Also, sustained
CC   release reduces the nature and severity of any side effects of the drug.
XX
SQ   Sequence   165 AA;

  Query Match             100.0%;  Score 846;  DB 22;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-86;
```

```
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 6
AAB66697
ID   AAB66697 standard; protein; 165 AA.
XX
AC   AAB66697;
XX
DT   06-APR-2001  (first entry)
XX
DE   Human erythropoietin protein #1.
XX
KW   Erythropoietin; EPO; reticulocytes; red blood cell;
KW   ethylene glycol; chronic renal failure; AIDS; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200102017-A2.
XX
PD   11-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-EP06009.
XX
PR   02-JUL-1999;   99US-0142243.
PR   05-AUG-1999;   99US-0147452.
PR   30-AUG-1999;   99US-0151454.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
PI   Burg J,  Hilger B,  Josel H;
XX
DR   WPI; 2001-147051/15.
XX
PT   Novel erythropoietin-glycoprotein conjugate useful for treating
PT   diseases correlated with anemia in chronic renal failure patients, AIDS
PT   and for treating cancer, is linked to polyethylene glycol through
PT   linker  -
XX
PS   Claim 19; Fig 1; 40pp; English.
XX
CC   The present invention relates to a conjugate comprising, human
CC   erythropoietin glycoprotein (EPO) having at least one free
CC   amino group and having in vivo biological activity of causing
CC   an increase the production of reticulocytes and red blood cells,
CC   covalently linked to 1-3 lower-alkoxy poly(ethylene glycol) groups
CC   through a linker. The invention is useful for preparation of
CC   medicaments for the treatment of prophylaxis of disease correlated
CC   with anemia in chronic renal failure patients (CRF), AIDS and for
CC   the treatment of cancer patients undergoing chemotherapy.
XX
SQ   Sequence  165 AA;

  Query Match           100.0%;  Score 846;  DB 22;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-86;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 7
ABB77896
ID   ABB77896 standard; protein; 165 AA.
XX
AC   ABB77896;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Amino acid sequence of a human erythropoietin (EPO).
XX
KW   Human; erythropoietin; EPO; glycoprotein; reticulocyte production;
KW   red blood cell production; anaemia; chronic renal failure;
KW   acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;
KW   committed erythroid progenitor.
XX
OS   Homo sapiens.
XX
PN   WO200249673-A2.
XX
PD   27-JUN-2002.
XX
PF   08-DEC-2001; 2001WO-EP14434.
XX
PR   20-DEC-2000; 2000EP-0127891.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
PI   Burg J,  Engel A,  Franze R,  Hilger B,  Schurig HE,  Tischer W;
PI   Wozny M;
XX
DR   WPI; 2002-566640/60.
XX
PT   Novel conjugate of erythropoietin glycoprotein with polyethylene
PT   glycol, useful for treating diseases correlated with anaemia in chronic
PT   renal failure patients and acquired immunodeficiency syndrome -
XX
PS   Claim 26; Fig 1; 40pp; English.
XX
CC   The present sequence represents a human erythropoietin (EPO) protein.
CC   It was used to produce conjugates of the invention. The specification
CC   describes a conjugate comprising an EPO glycoprotein having an N-terminal
CC   alpha-amino group, chosen from human EPO (hEPO) or its analogues (where
CC   hEPO is modified by addition of 1-6 glycosylation sites or a
CC   rearrangement of a glycosylation site). The glycoprotein is covalently
CC   linked to a poly(ethylene glycol) group. The EPO glycoprotein has in vivo
CC   biological activity of causing bone marrow cells to increase production
CC   of reticulocytes and red blood cells. The conjugate increased circulating
CC   half-life and plasma residence time, decreased clearance, increased
CC   clinical activity in vivo, improved potency and stability, when compared
CC   to unmodified EPO. The EPO conjugate is useful for preparing medicaments
CC   for the treatment and prophylaxis of diseases correlated with anaemia in
CC   chronic renal failure patients (CRF), acquired immunodeficiency syndrome
CC   (AIDS) and for treating cancer patients undergoing chemotherapy. It is
CC   also useful for treating patients by stimulating the division and
CC   differentiation of committed erythroid progenitors in the bone marrow.
XX
SQ   Sequence  165 AA;

  Query Match           100.0%;  Score 846;  DB 23;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-86;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 8
AAM53061
ID   AAM53061 standard; protein; 165 AA.
XX
AC   AAM53061;
XX
DT   25-MAR-2002  (first entry)
XX
DE   Human erythropoietin (hEPO), 165 residue form.
XX
KW   Human; erythropoietin; EPO; hEPO; haemostatic; red blood cell;
KW   blood disorder; anaemia; chronic renal failure; CRF; AIDS;
KW   acquired immunodeficiency syndrome; cancer chemotherapy; haemostatic;
KW   anti-HIV; antianaemic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Disulfide-bond  7..161
FT   Modified-site   24
FT                   /note= "N-glycosylated"
FT   Disulfide-bond  29..33
FT   Modified-site   38
FT                   /note= "N-glycosylated"
FT   Modified-site   83
FT                   /note= "N-glycosylated"
FT   Modified-site   126
FT                   /note= "O-glycosylated"
XX
PN   WO200187329-A1.
XX
PD   22-NOV-2001.
XX
PF   08-MAY-2001; 2001WO-EP05187.
XX
PR   15-MAY-2000; 2000EP-0110355.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
PI   Papadimitriou A;
XX
DR   WPI; 2002-082943/11.
XX
PT   Composition useful in the treatment of e.g. AIDS comprises an
PT   erythropoietin protein, and a multiple charged inorganic anion in a
PT   buffer
XX
PS   Claim 28; Fig 1; 64pp; English.
XX
CC   The invention relates to liquid pharmaceutical compositions comprising
CC   an erythropoietin (EPO) protein, a multiple negatively charged inorganic
CC   anion in a buffer which maintains the pH of the solution from 5.5-7.0,
CC   and optionally at least one excipient. The erythropoietin used in the
CC   composition is preferably human (AAM53061 or AAM53062) a human
CC   erythropoietin variant containing additional glycosylation sites
CC   (AAM53064-AAM53107), or an erythropoietin with the C-terminal addition
CC   of a C-terminal fragment of human chorionic gonadotropin (AAM53063).
CC   Erythropoietin is a glycoprotein essential for the formation of red blood
CC   cells and is therefore useful in the treatment of blood disorders
CC   characterised by low or defective red blood cell production. The
CC   compositions of the invention can be used in the treatment and prevention
CC   of anaemia in chronic renal failure patients (CRF), AIDS (acquired
CC   immunodeficiency syndrome), and/or for the treatment of cancer patients
CC   undergoing chemotherapy. Unlike prior art erythropoietin compositions,
CC   the compositions of the invention do not contain human serum albumin
CC   (thereby avoiding the possibility of viral infections and allergic
CC   reactions associated with this component), are liquid rather than
CC   lyophilisates (and therefore do not need to be reconstituted before
CC   administration), and are stable at elevated temperatures such as 25
CC   degrees Celsius and even 40 degrees Celsius, and therefore can be stored
CC   without refrigeration for prolonged periods without degradation and loss
CC   of activity. The present sequence represents the 165 residue form
CC   of human erythropoietin which is specifically claimed for use in a
CC   composition of the invention.
XX
SQ   Sequence   165 AA;

  Query Match             100.0%;  Score 846;  DB 23;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-86;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 9
AAP70398
ID   AAP70398 standard; protein; 166 AA.
XX
AC   AAP70398;
XX
DT   19-FEB-1991  (first entry)
XX
DE   Sequence of human erythropoietin (EPO).
XX
KW   Mega-karyocyte-platelet growth factor; hormone;
KW   mega-karyocyte colony stimulating factor; therapy;
KW   small acetyl cholinesterase positive cell;
KW   erythrocyte growth effect.
XX
OS   Homo sapiens.
XX
PN   JP62149624-A.
XX
PD   03-JUL-1987.
XX
PF   15-AUG-1986;   86JP-0191542.
XX
PR   13-SEP-1985;   85JP-0203049.
PR   15-AUG-1986;   86JP-0191542.
XX
PA   (KAWA/) KAWAKITA M.
XX
DR   WPI; 1987-224837/32.
XX
PT   Megakaryocyte-platelet growth factor - contains as active
PT   component human erythropoietin and is used to treat diseases
PT   caused by decrease in platelets
XX
PS   Disclosure; Page 181; 8pp; Japanese.
XX
CC   All of the Cys residues in the SQ are labelled "SH". Megakaryocyte-
CC   platelet growth factor contains human EPO as an active principle.
CC   Human EPO has a megakaryocyte colony-stimulating activity and
CC   increases the ratio of small acetyl cholinesterase posiitive cell
CC   (SAchE+) which is immature megakaryocyte. Human EPO effects
```

```
CC   megakaryocyte-platelet system other than an erythrocyte growth
CC   effect. Megakaryocyte-platelet growth is usable as a remedy for
CC   diseases caused by a platelet decrease.
XX
SQ   Sequence   166 AA;

  Query Match             100.0%;  Score 846;  DB 8;  Length 166;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-86;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 10
AAR23593
ID   AAR23593 standard; Protein; 166 AA.
XX
AC   AAR23593;
XX
DT   20-OCT-1992  (first entry)
XX
DE   Recombinant hematopoietic molecule portion 2.
XX
KW   Erythropoietin; EPO: erythrocytes; IL-3; haematopoiesis.
XX
OS   Homo sapiens.
XX
PN   WO9206116-A.
XX
PD   16-APR-1992.
XX
PF   26-SEP-1991;   91WO-US07053.
XX
PR   28-SEP-1990;   90US-0589958.
XX
PA   (ORTH ) ORTHO PHARM CORP.
XX
PI   Rosen JI;
XX
DR   WPI; 1992-150819/18.
XX
PT   Recombinant haematopoietic molecules useful in treating
PT   anaemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-5 or M-CSF
PT   and has early and later myeloid differentiation activity
XX
PS   Disclosure; Page 32; 82pp; English.
XX
CC   This protein sequence given comprises the entire amino acid sequence
CC   of human erythropoietin (EPO).  EPO leads to the maturation of
CC   erythrocytes and is therefore designated as a late myeloid
CC   differentiation factor (MDF).  Within the scope of the invention
CC   hybrid molecules were produced which contain at least a portion of an
CC   early MDF and at least a portion of a late MDF covalently linked.  The
CC   EPO sequence given is effective within the scope of the invention in
CC   full or in a truncated version.  Amino acids 7-161 act as a
CC   late MDF when recombined with an early MDF eg. IL-3.
CC   These compounds can be used to promote heamatopoiesis in a patient.
CC   The bonding of the early and late factors allows a very high conc. of
CC   late MDF at the surface of a cell which the early MDF is bound.  It
CC   also allows the early MDA to act more specifically to stimulate only
CC   the desired lineage, thus reducing undesirable effects.  These
CC   compounds are useful for treating anaemias of various origins eg.renal
CC   failure and AIDS.  It is easier to produce and administer one
CC   recombinant molecule rather than two separate molecules.
XX
SQ   Sequence   166 AA;

  Query Match             100.0%;  Score 846;  DB 13;  Length 166;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-86;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 11
AAW77780
ID   AAW77780 standard; Protein; 166 AA.
XX
AC   AAW77780;
XX
DT   24-NOV-1998  (first entry)
XX
DE   Human EPO receptor agonist polypeptide.
XX
KW   Haematopoietic receptor agonist; erythropoietin receptor agonist;
KW   EPO; human; chimeric protein; stem cell expansion; tumour;
KW   infection; autoimmune disease; haematopoietic disorder; therapy;
KW   dendritic cell.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..6
FT                   /note= "1-6 amino acids of the N-terminus are
FT                    optionally deleted"
FT   Misc-difference 162..166
FT                   /note= "1-5 amino acids of the C-terminus are
FT                    optionally deleted"
FT   Misc-difference 23..24
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 24..25
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 25..26
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 26..27
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 27..28
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 28..29
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 29..30
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 30..31
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 31..32
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 32..33
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 33..34
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 34..35
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 35..36
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 36..37
FT                   /note= "possible positions of new C- and N-termini"
```

```
FT   Misc-difference 37..38
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 38..39
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 39..40
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 40..41
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 41..42
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 42..43
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 43..44
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 44..45
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 45..46
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 46..47
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 47..48
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 48..49
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 49..50
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 50..51
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 51..52
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 52..53
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 53..54
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 54..55
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 55..56
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 56..57
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 57..58
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 77..78
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 78..79
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 79..80
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 81..82
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 82..83
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 84..85
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 85..86
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 86..87
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 87..88
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 88..89
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 108..109
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 109..110
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 110..111
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 111..112
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 112..113
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 113..114
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 114..115
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 115..116
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 116..117
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 117..118
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 118..119
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 119..120
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 120..121
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 121..122
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 122..123
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 123..124
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 124..125
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 125..126
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 126..127
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 127..128
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 128..129
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 129..130
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 130..131
FT                   /note= "possible positions of new C- and N-termini"
FT   Misc-difference 131..132
FT                   /note= "possible positions of new C- and N-termini"
XX
PN   WO9817810-A2.
XX
PD   30-APR-1998.
XX
PF   23-OCT-1997;   97WO-US20037.
XX
PR   25-OCT-1996;   96US-0029629.
XX
PA   (SEAR ) SEARLE & CO G D.
XX
PI   Feng Y,  McKearn JP,  McWherter CA,  Minnerly JC,  Minster NI;
PI   Staten NR,  Streeter PR,  Summers NL,  Woulfe SL;
XX
DR   WPI; 1998-261504/23.
XX
PT   Multi-functional chimeric haematopoietic receptor agonist - useful
PT   to treat haematopoietic disorders, tumours, infections or autoimmune
PT   diseases
XX
PS   Claim 1; Page 762; 841pp; English.
XX
CC   A human erythropoietin (EPO) receptor agonist polypeptide comprises
CC   a modified EPO amino acid sequence of the formula provided in
CC   AAW77780, in which the N-terminus is joined to the C-terminus directly
CC   or via a linker, the polypeptide having new C- and N-termini at one
CC   of the positions indicated.  Novel claimed multi-functional chimeric
CC   haematopoietic receptor agonists (see AAW77812-22) have the formula
CC   R1-L1-R2, R2-L1-R1, R1-R2 or R2-R1, where L is a linker and R1 and
CC   R2 are independently selected from: (a) the human EPO receptor
CC   agonist; (b) a human stem cell factor receptor agonist polypeptide
CC   (see AAW77781); (c) a human flt-3 receptor agonist polypeptide (see
CC   AAW77782); (d) a modified human granulocyte colony stimulating factor
CC   (G-CSF) polypeptide (see AAW77783); (e) modified human interleukin-3
CC   polypeptide (see AAW77784); (f) modified human c-mpl ligand polypeptide
CC   (see AAW77785); and (g) a factor selected from the group consisting of
```

CC a CSF, a cytokine, a lymphokine, an interleukin and a haematopoietic
CC growth factor, provided that at least R1 or R2 is selected from (a),
CC (b) or (c) as above. The multi-functional chimeric haematopoietic
CC receptor agonist can be used to stimulate the production of

Query Match 100.0%; Score 846; DB 19; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
Db 1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy 61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
Db 61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy 121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
Db 121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165

RESULT 12
AAW58404
ID AAW58404 standard; Protein; 166 AA.
XX
AC AAW58404;
XX
DT 12-OCT-1998 (first entry)
XX
DE Human erythropoietin.
XX
KW Erythropoietin receptor agonist; EPO; human; anaemia;
KW haematopoietic deficiency; red blood cell; erythroid progenitor;
KW bone marrow suppression.
XX
OS Homo sapiens.
XX
PN WO9818926-A1.
XX
PD 07-MAY-1998.
XX
PF 23-OCT-1997; 97WO-US18703.
XX
PR 25-OCT-1996; 96US-0034044.
XX
PA (SEAR) SEARLE & CO G D.
XX
PI Feng Y, McWherter CA, Summers N;
XX
DR WPI; 1998-272221/24.
DR N-PSDB; AAV31031.
XX
PT Human erythropoietin receptor agonist polypeptide - used to
PT stimulate the production of red blood cells in a patient
XX
PS Claim 1; Page 93; 112pp; English.
XX
CC A claimed human erythropoietin (EPO) receptor agonist polypeptide
CC comprises a modified EPO amino acid sequence given in AAW58404,
CC where (a) optionally 1-6 amino acids from the N-terminus and 1-5
CC from the C-terminus can be deleted, (b) the N-terminus is joined to
CC the C-terminus directly or through a linker (see AAW58405-12) capable
CC of joining the N-terminus to the C-terminus, (c) there are new C-
CC and N-termini at any two consecutive amino acids from amino acids
CC 23-24 to 38-39, 40-41 to 41-42, 43-44 to 48-49, 50-51 to 57-58,
CC 77-78 to 82-83, 84-85 to 88-89, and 108-109 to 131-132, and (d)
CC optionally the agonist polypeptide is preceded by Met, Ala, or
CC Met-Ala. 60 Of these circularly permuted EPO receptor agonists
CC (see AAW58413-72) are claimed. Also claimed are: nucleic acid
CC molecules (see AAV30971-V31030) encoding novel EPO receptor agonists;
CC a method of producing an EPO receptor agonist using transformed or
CC transfected host cells; and methods for stimulating the production
CC of haematopoietic cells, for selective ex vivo expansion of
CC erythroid progenitors, and treating patients having a haematopoietic
CC disorder using the EPO receptor agonists. The EPO receptor
CC agonists retain one or more activities of native EPO and may also
CC show improved haematopoietic cell-stimulating activity and/or an
CC improved activity profile which may include reduction of undesirable
CC biological activities associated with native EPO and/or have
CC improved physical properties such as increased solubility,
CC stability and refold efficiency.
XX
SQ Sequence 166 AA;

Query Match 100.0%; Score 846; DB 19; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
Db 1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy 61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
Db 61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy 121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
Db 121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165

RESULT 13
ABB07030
ID ABB07030 standard; Protein; 166 AA.
XX
AC ABB07030;
XX
DT 21-JUN-2002 (first entry)
XX
DE Modified erythropoietin related gene protein sequence.
XX
KW Modified erythropoietin; EPO.
XX
OS Unidentified.
XX
PN KR145802-B1.
XX
PD 01-AUG-1998.
XX
PF 31-MAY-1994; 94KR-0012082.
XX
PR 31-MAY-1994; 94KR-0012082.
XX
PA (GLDS) LG CHEM CO LTD.
XX
PI Kim C, Song Y, Lee T;
XX
DR WPI; 2000-234250/20.
DR N-PSDB; ABL50878.
XX
PT MODIFIED ERYTHROPOIETIN GENE AND EXPRESSION VECTORS THEREOF -
XX
PS Disclosure; Page 14; 15pp; Korean.
XX
CC The present invention describes modified erythropoietin (EPO) genes
CC and expression vectors comprising the genes. The present sequence
CC represents a protein sequence from the present invention.
XX
SQ Sequence 166 AA;

Query Match 100.0%; Score 846; DB 21; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
                 |||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 14
AAE02641
ID   AAE02641 standard; Protein; 166 AA.
XX
AC   AAE02641;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Human erythropoietin (EPO) mature protein.
XX
KW   Human; erythropoietin; EPO; antianaemic; nephrotrophic; anti-HIV;
KW   vaccine; haemostatic; immunoglobulin; Ig; EPO deficient disease;
KW   anaemia; renal failure; Human Immunodeficiency Virus; HIV;
KW   haematopoietic growth factor.
XX
OS   Homo sapiens.
XX
PN   WO200136489-A2.
XX
PD   25-MAY-2001.
XX
PF   03-NOV-2000; 2000WO-EP10843.
XX
PR   12-NOV-1999;   99US-0164855.
XX
PA   (MERE ) MERCK PATENT GMBH.
XX
PI   Hartmann A,  Brandt S,  Rieke E,  Sobel C,  Lo K,  Way JC,  Gillies S;
XX
DR   WPI; 2001-367563/38.
DR   N-PSDB; AAD06893.
XX
PT   Novel modified erythropoietin forms such as fusion proteins, comprising
PT   Fc portion of an immunoglobulin molecule and a target molecule having
PT   the biological activity of erythropoietin forms  -
XX
PS   Example 1; Page 22; 58pp; English.
XX
CC   The present sequence is human erythropoietin (EPO) mature protein. EPO
CC   has improved biological activity and an extended serum half life greater
CC   than 20 hours. The present invention relates to modified EPO forms such
CC   as fusion proteins comprising a Fc portion of an immunoglobulin (Ig)
CC   molecule and an EPO molecule (Fc-EPO). The Fc portion is fused covalently
CC   through its C-terminus directly or indirectly to the EPO molecule, and
CC   where the Fc portion as well as EPO portion may be modified or mutated.
CC   The invention also relates to non-fused EPO molecules which have a
CC   pattern of cysteines or disulphide bonding which is distinct from human
CC   or animal EPO. Pharmaceutical compositions containing EPO are useful in
CC   the treatment of EPO deficient diseases such as anaemia, renal failure,
CC   HIV infection, blood loss and chronic disease that can be treated with
CC   haematopoietic growth factor.
XX
SQ   Sequence   166 AA;

  Query Match             100.0%;  Score 846;  DB 22;  Length 166;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-86;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
                 |||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 15
AAB66698
ID   AAB66698 standard; protein; 166 AA.
XX
AC   AAB66698;
XX
DT   06-APR-2001  (first entry)
XX
DE   Human erythropoietin protein #2.
XX
KW   Erythropoietin; EPO; reticulocytes; red blood cell;
KW   ethylene glycol; chronic renal failure; AIDS; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200102017-A2.
XX
PD   11-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-EP06009.
XX
PR   02-JUL-1999;   99US-0142243.
PR   05-AUG-1999;   99US-0147452.
PR   30-AUG-1999;   99US-0151454.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
PI   Burg J,  Hilger B,  Josel H;
XX
DR   WPI; 2001-147051/15.
XX
PT   Novel erythropoietin-glycoprotein conjugate useful for treating
PT   diseases correlated with anemia in chronic renal failure patients, AIDS
PT   and for treating cancer, is linked to polyethylene glycol through
PT   linker  -
XX
PS   Claim 19; Fig 2; 40pp; English.
XX
CC   The present invention relates to a conjugate comprising, human
CC   erythropoietin glycoprotein (EPO) having at least one free
CC   amino group and having in vivo biological activity of causing
CC   an increase the production of reticulocytes and red blood cells,
CC   covalently linked to 1-3 lower-alkoxy poly(ethylene glycol) groups
CC   through a linker. The invention is useful for preparation of
CC   medicaments for the treatment of prophylaxis of disease correlated
CC   with anemia in chronic renal failure patients (CRF), AIDS and for
CC   the treatment of cancer patients undergoing chemotherapy.
XX
SQ   Sequence   166 AA;

  Query Match             100.0%;  Score 846;  DB 22;  Length 166;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-86;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
                 |||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
```

Search completed: June 13, 2003, 15:03:44
Job time : 40 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:03:00 ; Search time 22 Seconds
(without alignments)
220.672 Million cell updates/sec

Title: US-09-830-964-1
Perfect score: 846
Sequence: 1 APPRLICDSRVLERYLLEAK.....SNFLRGKLKLYTGEACRTGD 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	846	100.0	165	4	US-09-604-871-1	Sequence 1, Appli
2	846	100.0	166	1	US-08-318-193-70	Sequence 70, Appl
3	846	100.0	166	4	US-09-604-871-2	Sequence 2, Appli
4	846	100.0	166	5	PCT-US94-04361-37	Sequence 37, Appl
5	846	100.0	193	1	US-07-903-220-1	Sequence 1, Appli
6	846	100.0	193	2	US-08-883-795A-34	Sequence 34, Appl
7	843	99.6	412	4	US-09-366-009-34	Sequence 34, Appl
8	830	98.1	166	5	PCT-US94-04361-45	Sequence 45, Appl
9	764.5	90.4	165	5	PCT-US94-04361-38	Sequence 38, Appl
10	759.5	89.8	165	5	PCT-US94-04361-39	Sequence 39, Appl
11	713	84.3	166	5	PCT-US94-04361-44	Sequence 44, Appl
12	701	82.9	166	5	PCT-US94-04361-41	Sequence 41, Appl
13	689	81.4	166	5	PCT-US94-04361-40	Sequence 40, Appl
14	678.5	80.2	167	5	PCT-US94-04361-42	Sequence 42, Appl
15	678	80.1	168	5	PCT-US94-04361-43	Sequence 43, Appl
16	193	22.8	36	5	PCT-US94-04361-50	Sequence 50, Appl
17	128	15.1	27	5	PCT-US94-04361-51	Sequence 51, Appl
18	114.5	13.5	34	5	PCT-US94-04361-52	Sequence 52, Appl
19	107	12.6	21	5	PCT-US94-04361-49	Sequence 49, Appl
20	103	12.2	20	2	US-08-759-599-9	Sequence 9, Appli
21	103	12.2	20	4	US-09-294-457-9	Sequence 9, Appli
22	101	11.9	20	5	PCT-US94-04361-58	Sequence 58, Appl
23	92	10.9	332	4	US-08-875-533-66	Sequence 66, Appl
24	91	10.8	17	1	US-08-232-513A-11	Sequence 11, Appl
25	91	10.8	17	4	US-09-231-159-16	Sequence 16, Appl
26	91	10.8	17	4	US-08-611-307-16	Sequence 16, Appl
27	91	10.8	312	4	US-08-875-533-67	Sequence 67, Appl
28	89	10.5	153	4	US-08-875-533-65	Sequence 65, Appl
29	89	10.5	165	1	US-08-413-803-29	Sequence 29, Appl
30	89	10.5	174	3	US-08-471-045-56	Sequence 56, Appl
31	89	10.5	174	3	US-08-469-712A-56	Sequence 56, Appl
32	89	10.5	174	4	US-08-446-871-56	Sequence 56, Appl
33	89	10.5	174	4	US-08-468-910-56	Sequence 56, Appl
34	89	10.5	174	4	US-08-761-907-56	Sequence 56, Appl
35	89	10.5	174	5	PCT-US95-03776-27	Sequence 27, Appl
36	89	10.5	195	1	US-08-388-779A-4	Sequence 4, Appli
37	89	10.5	195	1	US-08-591-070A-4	Sequence 4, Appli
38	89	10.5	195	2	US-08-927-855-4	Sequence 4, Appli
39	89	10.5	285	4	US-08-875-533-69	Sequence 69, Appl
40	89	10.5	285	4	US-08-875-533-70	Sequence 70, Appl
41	89	10.5	288	4	US-08-875-533-71	Sequence 71, Appl
42	89	10.5	288	4	US-08-875-533-73	Sequence 73, Appl
43	89	10.5	332	4	US-08-875-533-22	Sequence 22, Appl
44	89	10.5	332	5	PCT-US95-03776-25	Sequence 25, Appl
45	89	10.5	353	1	US-08-330-517-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-604-871-1
; Sequence 1, Application US/09604871
; Patent No. 6340742
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Hilger, Bernd
; APPLICANT: Josel, Hans-Peter
; TITLE OF INVENTION: ERYTHROPOIETIN CONJUGATES
; FILE REFERENCE: 1098 nonprovisional
; CURRENT APPLICATION NUMBER: US/09/604,871
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/151,454
; PRIOR FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: 60/147,452
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/142,243
; PRIOR FILING DATE: 1999-07-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-604-871-1

Query Match 100.0%; Score 846; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 5.3e-100;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
```

RESULT 2
US-08-318-193-70
; Sequence 70, Application US/08318193
; Patent No. 5641663
; GENERAL INFORMATION:
; APPLICANT: GARVIN, Robert T.
; APPLICANT: MALEK, Lawrence T.

```
;    TITLE OF INVENTION:  AN EXPRESSION SYSTEM FOR THE SECRETION
;    TITLE OF INVENTION:  OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY
;    TITLE OF INVENTION:  STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS
;    TITLE OF INVENTION:  PROTEINS FROM STREPTOMYCES
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/318,193
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,314
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/224,568
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  18740/116 CACO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  70:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-318-193-70

  Query Match             100.0%;  Score 846;  DB 1;  Length 166;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-100;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 3
US-09-604-871-2
; Sequence 2, Application US/09604871
; Patent No. 6340742
; GENERAL INFORMATION:
;  APPLICANT:  Burg, Josef
;  APPLICANT:  Hilger, Bernd
;  APPLICANT:  Josel, Hans-Peter
;  TITLE OF INVENTION:  ERYTHROPOIETIN CONJUGATES
;  FILE REFERENCE:  1098 nonprovisional
;  CURRENT APPLICATION NUMBER: US/09/604,871
;  CURRENT FILING DATE:  2000-06-28
;  PRIOR APPLICATION NUMBER:  60/151,454
;  PRIOR FILING DATE:  1999-08-30
;  PRIOR APPLICATION NUMBER:  60/147,452
;  PRIOR FILING DATE:  1999-08-05
;  PRIOR APPLICATION NUMBER:  60/142,243
;  PRIOR FILING DATE:  1999-07-02
;  NUMBER OF SEQ ID NOS:  3
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-604-871-2

  Query Match             100.0%;  Score 846;  DB 4;  Length 166;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-100;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 4
PCT-US94-04361-37
; Sequence 37, Application PC/TUS9404361
;  GENERAL INFORMATION:
;    APPLICANT:  Brigham and Women's Hospital
;    APPLICANT:  75 Francis Street
;    APPLICANT:  Boston, MA 02115
;    APPLICANT:  Bunn, H. Franklin
;    APPLICANT:  Wen, Danyi
;    APPLICANT:  Showers, Mark O.
;    TITLE OF INVENTION:  Erythropoietin Muteins With Enhanced
;    TITLE OF INVENTION:  Activity
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04361
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/049,802
;      FILING DATE:  21-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cimbala, Michele A.
;      REGISTRATION NUMBER:  33,851
;      REFERENCE/DOCKET NUMBER:  0627.336PC01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  both
PCT-US94-04361-37
```

```
  Query Match             100.0%;  Score 846;  DB 5;  Length 166;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-100;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 5
US-07-903-220-1
; Sequence 1, Application US/07903220
; Patent No. 5322837
;  GENERAL INFORMATION:
;    APPLICANT:  Hewick, Rodney M.
;    TITLE OF INVENTION:  METHOD FOR THE PURIFICATION OF
;    TITLE OF INVENTION:  ERYTHROPOIETIN AND ERYTHROPOIETIN COMPOSITION
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Paul H. Heller
;      STREET:  Kenyon & Kenyon, One Broadway
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/903,220
;      FILING DATE:  19920731
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  1248/27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 429-1776
;      TELEFAX:  (202) 429-0796
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  193 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
US-07-903-220-1

  Query Match             100.0%;  Score 846;  DB 1;  Length 193;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-100;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 147

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        148 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 192


RESULT 6
US-08-883-795A-34
; Sequence 34, Application US/08883795A
; Patent No. 5985607
;  GENERAL INFORMATION:
;    APPLICANT:  Delcuve, Genevieve
;    APPLICANT:  Awang, Gregor
;    TITLE OF INVENTION:  Recombinant DNA Molecules and Expression
;    TITLE OF INVENTION:  Vectors for Tissue Plasminogen Activator
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BERESKIN & PARR
;      STREET:  40 King Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/883,795A
;      FILING DATE:  27-JUN-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gravelle, Micheline
;      REGISTRATION NUMBER:  40,261
;      REFERENCE/DOCKET NUMBER:  7841-062
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  193 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-883-795A-34

  Query Match             100.0%;  Score 846;  DB 2;  Length 193;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-100;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 147

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        148 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 192


RESULT 7
US-09-366-009-34
; Sequence 34, Application US/09366009
; Patent No. 6426042
;  GENERAL INFORMATION:
;       APPLICANT: Asada, Kiyozo
;                  Uemori, Takashi
;                  Ueno, Takashi
;                  Koyama, No. 6426042uto
;                  Hashino, Kimikazu
```

```
;                Kato, Ikunoshin
;        TITLE OF INVENTION: METHOD FOR GENE TRANSFER INTO TARGET
;                            CELLS WITH RETROVIRUS
;        NUMBER OF SEQUENCES: 39
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: WEISER & ASSOCIATES
;             STREET: 230 South Fifteenth Street, Suite 500
;             CITY: Philadelphia
;             STATE: PA
;             COUNTRY: USA
;             ZIP: 19102
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/366,009
;             FILING DATE: 02-Aug-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/809,156
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: JP 294382/1995
;             FILING DATE: 13-NOV-1995
;             APPLICATION NUMBER: JP 051847/1996
;             FILING DATE: 08-MAR-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Weiser, Gerard J.
;             REGISTRATION NUMBER: 19,763
;             REFERENCE/DOCKET NUMBER: 977.6507P
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 215-875-8383
;             TELEFAX: 215-875-8394
;    INFORMATION FOR SEQ ID NO: 34:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 412 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-366-009-34

  Query Match             99.6%;  Score 843;  DB 4;  Length 412;
  Best Local Similarity   99.4%;  Pred. No. 5.5e-99;
  Matches  164;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 APPRLICDSRVLQRYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 292

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 352

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        353 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 397


RESULT 8
PCT-US94-04361-45
; Sequence 45, Application PC/TUS9404361
;  GENERAL INFORMATION:
;    APPLICANT:  Brigham and Women's Hospital
;    APPLICANT:  75 Francis Street
;    APPLICANT:  Boston, MA 02115
;    APPLICANT:  Bunn, H. Franklin
;    APPLICANT:  Wen, Danyi
;    APPLICANT:  Showers, Mark O.
;    TITLE OF INVENTION:  Erythropoietin Muteins With Enhanced
;    TITLE OF INVENTION:  Activity
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04361
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/049,802
;      FILING DATE:  21-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cimbala, Michele A.
;      REGISTRATION NUMBER:  33,851
;      REFERENCE/DOCKET NUMBER:  0627.336PC01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  both
PCT-US94-04361-45

  Query Match             98.1%;  Score 830;  DB 5;  Length 166;
  Best Local Similarity   98.8%;  Pred. No. 5.9e-98;
  Matches  163;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEGVQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 9
PCT-US94-04361-38
; Sequence 38, Application PC/TUS9404361
;  GENERAL INFORMATION:
;    APPLICANT:  Brigham and Women's Hospital
;    APPLICANT:  75 Francis Street
;    APPLICANT:  Boston, MA 02115
;    APPLICANT:  Bunn, H. Franklin
;    APPLICANT:  Wen, Danyi
;    APPLICANT:  Showers, Mark O.
;    TITLE OF INVENTION:  Erythropoietin Muteins With Enhanced
;    TITLE OF INVENTION:  Activity
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04361
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/049,802
;      FILING DATE:  21-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cimbala, Michele A.
;      REGISTRATION NUMBER:  33,851
;      REFERENCE/DOCKET NUMBER:  0627.336PC01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  165 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  both
PCT-US94-04361-38

  Query Match             90.4%;  Score 764.5;  DB 5;  Length 165;
  Best Local Similarity   91.5%;  Pred. No. 1.4e-89;
  Matches  151;  Conservative    7;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||||||||||||||||||||:| ||:| ||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENVTMGCSESCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||:| ||||:||||||:|||:||||||:|||||||||| ||||
Db         61 VEVWQGLALLSEAVLRGQAVLANSSQPFEPLQLHMDKAISGLRSITTLLRALGAQ-EAIS 119

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              ||||||||||||||||| |||||||||||||||||||| ||
Db        120 LPDAASAAPLRTITADTFCKLFRVYSNFLRGKLKLYTGEACRRGD 164


RESULT 10
PCT-US94-04361-39
; Sequence 39, Application PC/TUS9404361
;  GENERAL INFORMATION:
;    APPLICANT:  Brigham and Women's Hospital
;    APPLICANT:  75 Francis Street
;    APPLICANT:  Boston, MA 02115
;    APPLICANT:  Bunn, H. Franklin
;    APPLICANT:  Wen, Danyi
;    APPLICANT:  Showers, Mark O.
;    TITLE OF INVENTION:  Erythropoietin Muteins With Enhanced
;    TITLE OF INVENTION:  Activity
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04361
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/049,802
;      FILING DATE:  21-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cimbala, Michele A.
;      REGISTRATION NUMBER:  33,851
;      REFERENCE/DOCKET NUMBER:  0627.336PC01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  165 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  both
PCT-US94-04361-39

  Query Match             89.8%;  Score 759.5;  DB 5;  Length 165;
  Best Local Similarity   90.3%;  Pred. No. 5.9e-89;
  Matches  149;  Conservative    9;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||:||||||||||||||||:| ||:| ||||||||||||||||||||:||||||
Db          1 APPRLVCDSRVLERYLLEAKEAENVTMGCSESCSLNENITVPDTKVNFYAWKRIEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||:| |||||:||||||:|||:||||||:|||||||||| ||||
Db         61 VEVWQGLALLSEAVLRGQAVLANSSQPFEPLQLHMDKAISGLRSITTLLRALGAQ-EAIS 119

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              ||||||||||||||||| |||||||||||||||||||| ||
Db        120 LPDAASAAPLRTITADTFCKLFRVYSNFLRGKLKLYTGEACRRGD 164


RESULT 11
PCT-US94-04361-44
; Sequence 44, Application PC/TUS9404361
;  GENERAL INFORMATION:
;    APPLICANT:  Brigham and Women's Hospital
;    APPLICANT:  75 Francis Street
;    APPLICANT:  Boston, MA 02115
;    APPLICANT:  Bunn, H. Franklin
;    APPLICANT:  Wen, Danyi
;    APPLICANT:  Showers, Mark O.
;    TITLE OF INVENTION:  Erythropoietin Muteins With Enhanced
;    TITLE OF INVENTION:  Activity
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04361
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/049,802
;      FILING DATE:  21-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cimbala, Michele A.
;      REGISTRATION NUMBER:  33,851
;      REFERENCE/DOCKET NUMBER:  0627.336PC01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
```

```
;      TYPE:  amino acid
;      TOPOLOGY:  both
PCT-US94-04361-44

  Query Match             84.3%;  Score 713;  DB 5;  Length 166;
  Best Local Similarity   84.2%;  Pred. No. 5.2e-83;
  Matches  139;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||||||||||||:|||:||||:| |||| || :|||||||||||| ||||:|||||
Db          1 APPRLICDSRVLERYILEAREAENVTMGCAEGCSFSENITVPDTKVNFYTWKRMDVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||:|||||| ||||| | |||||||||| |||||:|||||||||| |
Db         61 VEVWQGLALLSEAILRGQALLANSSQPSETLQLHVDKAVSSLRSLTSLLRALGAQKEATS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
                |:| ||||||| | ||  ||||:||||||||| ||||||||  ||
Db        121 LPEATSAAPLRTFTVDTLCKLFRIYSNFLRGKLTLYTGEACRRGD 165


RESULT 12
PCT-US94-04361-41
; Sequence 41, Application PC/TUS9404361
;  GENERAL INFORMATION:
;    APPLICANT:  Brigham and Women's Hospital
;    APPLICANT:  75 Francis Street
;    APPLICANT:  Boston, MA 02115
;    APPLICANT:  Bunn, H. Franklin
;    APPLICANT:  Wen, Danyi
;    APPLICANT:  Showers, Mark O.
;    TITLE OF INVENTION:  Erythropoietin Muteins With Enhanced
;    TITLE OF INVENTION:  Activity
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04361
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/049,802
;      FILING DATE:  21-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cimbala, Michele A.
;      REGISTRATION NUMBER:  33,851
;      REFERENCE/DOCKET NUMBER:  0627.336PC01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  both
PCT-US94-04361-41

  Query Match             82.9%;  Score 701;  DB 5;  Length 166;
  Best Local Similarity   82.4%;  Pred. No. 1.8e-81;
  Matches  136;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||||||||||||:|||||||:| ||||   |:||||||||||||||||||:| :||

Db          1 APPRLICDSRVLERYILEAKEAENVTMGCAEGPRLSENITVPDTKVNFYAWKRMKVEEQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              |||||||:|||||:|: |||  ||||| | ||||:|||:|||||||:||| |||||| :|
Db         61 VEVWQGLSLLSEAILQAQALQANSSQPPESLQLHIDKAISGLRSLTSLLRVLGAQKELMS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
                ||||  |||||:|||| ||||||||||||||||||||||  ||
Db        121 PPDATQAAPLRTLTADTFCKLFRVYSNFLRGKLKLYTGEACRRGD 165


RESULT 13
PCT-US94-04361-40
; Sequence 40, Application PC/TUS9404361
;  GENERAL INFORMATION:
;    APPLICANT:  Brigham and Women's Hospital
;    APPLICANT:  75 Francis Street
;    APPLICANT:  Boston, MA 02115
;    APPLICANT:  Bunn, H. Franklin
;    APPLICANT:  Wen, Danyi
;    APPLICANT:  Showers, Mark O.
;    TITLE OF INVENTION:  Erythropoietin Muteins With Enhanced
;    TITLE OF INVENTION:  Activity
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04361
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/049,802
;      FILING DATE:  21-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cimbala, Michele A.
;      REGISTRATION NUMBER:  33,851
;      REFERENCE/DOCKET NUMBER:  0627.336PC01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  both
PCT-US94-04361-40

  Query Match             81.4%;  Score 689;  DB 5;  Length 166;
  Best Local Similarity   80.0%;  Pred. No. 6.1e-80;
  Matches  132;  Conservative   14;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||||||||||||:|||||||:| ||||   |:||||||||||||||||||:| :||
Db          1 APPRLICDSRVLERYILEAKEAENVTMGCAEGPRLSENITVPDTKVNFYAWKRMEVEEQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :||||||:|||||:|: |||| ||||| | ||||:|||:|||||||:||| |||||| :|
Db         61 IEVWQGLSLLSEAILQAQALLANSSQPPETLQLHIDKAISGLRSLTSLLRVLGAQKELMS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
                |||    |||||:| ||| ||||||:|||||||||||||| || ||
Db        121 PPDTTPPAPLRTLTVDTFCKLFRVYANFLRGKLKLYTGEVCRRGD 165
```

```
RESULT 14
PCT-US94-04361-42
; Sequence 42, Application PC/TUS9404361
;  GENERAL INFORMATION:
;     APPLICANT:  Brigham and Women's Hospital
;     APPLICANT:  75 Francis Street
;     APPLICANT:  Boston, MA 02115
;     APPLICANT:  Bunn, H. Franklin
;     APPLICANT:  Wen, Danyi
;     APPLICANT:  Showers, Mark O.
;     TITLE OF INVENTION:  Erythropoietin Muteins With Enhanced
;     TITLE OF INVENTION:  Activity
;     NUMBER OF SEQUENCES:  59
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;       STREET:  1100 New York Avenue, Suite 600
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20005-3934
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/04361
;       FILING DATE:  Herewith
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/049,802
;       FILING DATE:  21-APR-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Cimbala, Michele A.
;       REGISTRATION NUMBER:  33,851
;       REFERENCE/DOCKET NUMBER:  0627.336PC01
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 371-2600
;       TELEFAX:  (202) 371-2540
;   INFORMATION FOR SEQ ID NO:  42:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  167 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  both
PCT-US94-04361-42

  Query Match             80.2%;  Score 678.5;  DB 5;  Length 167;
  Best Local Similarity   81.3%;  Pred. No. 1.3e-78;
  Matches  135;  Conservative    9;  Mismatches   21;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||:|||:|||| | |||| || :||||||||||||||||||| |||
Db          1 APPRLICDSRVLERYILEAREAENATMGCAEGCSFSENITVPDTKVNFYAWKRMEVQQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :|||||||||||||: ||||| |:||| | |:|||||||||||||:||||||||||||
Db         61 LEVWQGLALLSEAIFRGQALPANASQPCEALRLHVDKAVSGLRSLTSLLRALGAQKEAIP 120

Qy        121 PPDAA-SAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
               |||  |||||  | |   ||||:|||||||| ||||||| ||
Db        121 LPDATPSAAPLRIFTVDALSKLFRIYSNFLRGKLTLYTGEACRRGD 166


RESULT 15
PCT-US94-04361-43
; Sequence 43, Application PC/TUS9404361
;  GENERAL INFORMATION:
;     APPLICANT:  Brigham and Women's Hospital
;     APPLICANT:  75 Francis Street
;     APPLICANT:  Boston, MA 02115
;     APPLICANT:  Bunn, H. Franklin
;     APPLICANT:  Wen, Danyi
;     APPLICANT:  Showers, Mark O.
;     TITLE OF INVENTION:  Erythropoietin Muteins With Enhanced
;     TITLE OF INVENTION:  Activity
;     NUMBER OF SEQUENCES:  59
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;       STREET:  1100 New York Avenue, Suite 600
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20005-3934
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/04361
;       FILING DATE:  Herewith
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/049,802
;       FILING DATE:  21-APR-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Cimbala, Michele A.
;       REGISTRATION NUMBER:  33,851
;       REFERENCE/DOCKET NUMBER:  0627.336PC01
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 371-2600
;       TELEFAX:  (202) 371-2540
;   INFORMATION FOR SEQ ID NO:  43:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  168 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  both
PCT-US94-04361-43

  Query Match             80.1%;  Score 678;  DB 5;  Length 168;
  Best Local Similarity   82.0%;  Pred. No. 1.6e-78;
  Matches  137;  Conservative    7;  Mismatches   21;  Indels    2;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||:|||| || | |||| || :||||||||||||||||||| |||
Db          1 APPRLICDSRVLERYILEAKEGENATMGCAESCSFSENITVPDTKVNFYAWKRMEVQQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :|||||||||||||:|:|||| |||| | ||||||||||||||||:||||||||||
Db         61 MEVWQGLALLSEAILQGQALLANSSQPSEALQLHVDKAVSGLRSLTSLLRALGAQKEAIP 120

Qy        121 PPDA--ASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
               |||  :|| ||||   ||  |||| |||||||| ||||||||  |
Db        121 LPDASPSSATPLRTFAVDTLCKLFRNYSNFLRGKLTLYTGEACRRRD 167


Search completed: June 13, 2003, 15:05:04
Job time : 24 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:04:47 ; Search time 47 Seconds
(without alignments)
375.342 Million cell updates/sec

Title: US-09-830-964-1
Perfect score: 846
Sequence: 1 APPRLICDSRVLERYLLEAK.....SNFLRGKLKLYTGEACRTGD 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	846	100.0	165	9	US-10-241-356-1	Sequence 1, Appli
2	846	100.0	165	9	US-09-945-517-1	Sequence 1, Appli
3	846	100.0	165	10	US-09-853-731-1	Sequence 1, Appli
4	846	100.0	165	12	US-10-014-363-1	Sequence 1, Appli
5	846	100.0	166	9	US-10-241-356-2	Sequence 2, Appli
6	846	100.0	166	10	US-09-853-731-2	Sequence 2, Appli
7	846	100.0	166	12	US-10-014-363-2	Sequence 2, Appli
8	846	100.0	169	12	US-10-014-363-4	Sequence 4, Appli
9	846	100.0	174	12	US-10-014-363-3	Sequence 3, Appli
10	846	100.0	174	12	US-10-014-363-5	Sequence 5, Appli
11	846	100.0	193	9	US-10-113-824-2	Sequence 2, Appli
12	846	100.0	193	9	US-09-813-775C-4	Sequence 4, Appli
13	846	100.0	435	9	US-09-932-812-22	Sequence 22, Appl
14	846	100.0	436	9	US-09-932-812-18	Sequence 18, Appl
15	846	100.0	437	9	US-09-932-812-20	Sequence 20, Appl
16	843	99.6	412	9	US-09-775-964-34	Sequence 34, Appl
17	838	99.1	193	9	US-09-813-775C-2	Sequence 2, Appli
18	834	98.6	193	9	US-09-813-775C-5	Sequence 5, Appli
19	825	97.5	166	9	US-09-813-775C-30	Sequence 30, Appl
20	825	97.5	193	9	US-09-813-775C-46	Sequence 46, Appl
21	824	97.4	166	9	US-09-813-775C-22	Sequence 22, Appl
22	824	97.4	166	9	US-09-813-775C-32	Sequence 32, Appl
23	824	97.4	193	9	US-09-813-775C-38	Sequence 38, Appl
24	824	97.4	193	9	US-09-813-775C-48	Sequence 48, Appl
25	822	97.2	166	9	US-09-813-775C-20	Sequence 20, Appl
26	822	97.2	166	9	US-09-813-775C-24	Sequence 24, Appl
27	822	97.2	193	9	US-09-813-775C-36	Sequence 36, Appl
28	822	97.2	193	9	US-09-813-775C-40	Sequence 40, Appl
29	821	97.0	166	9	US-09-813-775C-26	Sequence 26, Appl
30	821	97.0	166	9	US-09-813-775C-31	Sequence 31, Appl
31	821	97.0	193	9	US-09-813-775C-42	Sequence 42, Appl
32	821	97.0	193	9	US-09-813-775C-47	Sequence 47, Appl
33	820	96.9	166	9	US-09-813-775C-18	Sequence 18, Appl
34	820	96.9	166	9	US-09-813-775C-23	Sequence 23, Appl
35	820	96.9	166	9	US-09-813-775C-28	Sequence 28, Appl
36	820	96.9	166	9	US-09-813-775C-33	Sequence 33, Appl
37	820	96.9	193	9	US-09-813-775C-34	Sequence 34, Appl
38	820	96.9	193	9	US-09-813-775C-39	Sequence 39, Appl
39	820	96.9	193	9	US-09-813-775C-44	Sequence 44, Appl
40	820	96.9	193	9	US-09-813-775C-49	Sequence 49, Appl
41	818	96.7	166	9	US-09-813-775C-21	Sequence 21, Appl
42	818	96.7	166	9	US-09-813-775C-25	Sequence 25, Appl
43	818	96.7	193	9	US-09-813-775C-37	Sequence 37, Appl
44	818	96.7	193	9	US-09-813-775C-41	Sequence 41, Appl
45	817	96.6	166	9	US-09-813-775C-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-10-241-356-1
; Sequence 1, Application US/10241356
; Publication No. US20030077753A1
; GENERAL INFORMATION:
; APPLICANT: TISCHER, WILHELM
; TITLE OF INVENTION: DIGLYCOSYLATED ERYTHROPOIETIN
; FILE REFERENCE: 20971
; CURRENT APPLICATION NUMBER: US/10/241,356
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: EP 01122555.4
; PRIOR FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-241-356-1

Query Match 100.0%; Score 846; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 8.7e-84;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
```

RESULT 2
US-09-945-517-1
; Sequence 1, Application US/09945517
; Publication No. US20030104996A1
; GENERAL INFORMATION:

```
;  APPLICANT: Li, Tiansheng
;  APPLICANT:  Chang, Byeong
;  APPLICANT:  Sloey, Christopher
;  TITLE OF INVENTION: L-METHIONINE AS A STABILIZER FOR NESP/EPO IN HSA-FREE FORMULATION
;  FILE REFERENCE: A-803
;  CURRENT APPLICATION NUMBER: US/09/945,517
;  CURRENT FILING DATE:  2001-08-30
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-945-517-1

  Query Match           100.0%;  Score 846;  DB 9;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 8.7e-84;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 3
US-09-853-731-1
; Sequence 1, Application US/09853731
; Patent No. US20020037841A1
; GENERAL INFORMATION:
;  APPLICANT: Papadimitriou, Apollon
;  TITLE OF INVENTION: Erythropoietin Composition
;  FILE REFERENCE: 20619 US
;  CURRENT APPLICATION NUMBER: US/09/853,731
;  CURRENT FILING DATE:  2001-05-11
;  PRIOR APPLICATION NUMBER: EP/00110355.5
;  PRIOR FILING DATE: 2000-05-15
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-853-731-1

  Query Match           100.0%;  Score 846;  DB 10;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 8.7e-84;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 4
US-10-014-363-1
; Sequence 1, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
;  APPLICANT: Burg, Josef
;  APPLICANT:  Engel, Alfred
;  APPLICANT:  Franze, Reinhard
;  APPLICANT:  Hilger, Bernd
;  APPLICANT:  Schurig, Hartmut Ernst
;  APPLICANT:  Tischer, Wilhelm
;  APPLICANT:  Wozny, Manfred
;  TITLE OF INVENTION: Erythropoietin Conjugates
;  FILE REFERENCE: Case 20805
;  CURRENT APPLICATION NUMBER: US/10/014,363
;  CURRENT FILING DATE:  2001-12-11
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-014-363-1

  Query Match           100.0%;  Score 846;  DB 12;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 8.7e-84;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 5
US-10-241-356-2
; Sequence 2, Application US/10241356
; Publication No. US20030077753A1
; GENERAL INFORMATION:
;  APPLICANT: TISCHER, WILHELM
;  TITLE OF INVENTION: DIGLYCOSYLATED ERYTHROPOIETIN
;  FILE REFERENCE: 20971
;  CURRENT APPLICATION NUMBER: US/10/241,356
;  CURRENT FILING DATE:  2002-09-11
;  PRIOR APPLICATION NUMBER: EP 01122555.4
;  PRIOR FILING DATE: 2001-09-25
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-241-356-2

  Query Match           100.0%;  Score 846;  DB 9;  Length 166;
  Best Local Similarity  100.0%;  Pred. No. 8.7e-84;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
```

```
RESULT 6
US-09-853-731-2
; Sequence 2, Application US/09853731
; Patent No. US20020037841A1
; GENERAL INFORMATION:
;  APPLICANT: Papadimitriou, Apollon
;  TITLE OF INVENTION: Erythropoietin Composition
;  FILE REFERENCE: 20619 US
;  CURRENT APPLICATION NUMBER: US/09/853,731
;  CURRENT FILING DATE:  2001-05-11
;  PRIOR APPLICATION NUMBER: EP/00110355.5
;  PRIOR FILING DATE: 2000-05-15
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-853-731-2

  Query Match            100.0%;  Score 846;  DB 10;  Length 166;
  Best Local Similarity  100.0%;  Pred. No. 8.7e-84;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 7
US-10-014-363-2
; Sequence 2, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
;  APPLICANT: Burg, Josef
;  APPLICANT:  Engel, Alfred
;  APPLICANT:  Franze, Reinhard
;  APPLICANT:  Hilger, Bernd
;  APPLICANT:  Schurig, Hartmut Ernst
;  APPLICANT:  Tischer, Wilhelm
;  APPLICANT:  Wozny, Manfred
;  TITLE OF INVENTION: Erythropoietin Conjugates
;  FILE REFERENCE: Case 20805
;  CURRENT APPLICATION NUMBER: US/10/014,363
;  CURRENT FILING DATE:  2001-12-11
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-014-363-2

  Query Match            100.0%;  Score 846;  DB 12;  Length 166;
  Best Local Similarity  100.0%;  Pred. No. 8.7e-84;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 8
US-10-014-363-4
; Sequence 4, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
;  APPLICANT: Burg, Josef
;  APPLICANT:  Engel, Alfred
;  APPLICANT:  Franze, Reinhard
;  APPLICANT:  Hilger, Bernd
;  APPLICANT:  Schurig, Hartmut Ernst
;  APPLICANT:  Tischer, Wilhelm
;  APPLICANT:  Wozny, Manfred
;  TITLE OF INVENTION: Erythropoietin Conjugates
;  FILE REFERENCE: Case 20805
;  CURRENT APPLICATION NUMBER: US/10/014,363
;  CURRENT FILING DATE:  2001-12-11
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 169
;   TYPE: PRT
;   ORGANISM: CHO/dhfr-
US-10-014-363-4

  Query Match            100.0%;  Score 846;  DB 12;  Length 169;
  Best Local Similarity  100.0%;  Pred. No. 9e-84;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 63

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 123

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        124 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 168


RESULT 9
US-10-014-363-3
; Sequence 3, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
;  APPLICANT: Burg, Josef
;  APPLICANT:  Engel, Alfred
;  APPLICANT:  Franze, Reinhard
;  APPLICANT:  Hilger, Bernd
;  APPLICANT:  Schurig, Hartmut Ernst
;  APPLICANT:  Tischer, Wilhelm
;  APPLICANT:  Wozny, Manfred
;  TITLE OF INVENTION: Erythropoietin Conjugates
;  FILE REFERENCE: Case 20805
;  CURRENT APPLICATION NUMBER: US/10/014,363
;  CURRENT FILING DATE:  2001-12-11
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 174
;   TYPE: PRT
;   ORGANISM: CHO/dhfr-
US-10-014-363-3

  Query Match            100.0%;  Score 846;  DB 12;  Length 174;
  Best Local Similarity  100.0%;  Pred. No. 9.3e-84;
```

```
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 68

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 128

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        129 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 173


RESULT 10
US-10-014-363-5
; Sequence 5, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
;  APPLICANT: Burg, Josef
;  APPLICANT:  Engel, Alfred
;  APPLICANT:  Franze, Reinhard
;  APPLICANT:  Hilger, Bernd
;  APPLICANT:  Schurig, Hartmut Ernst
;  APPLICANT:  Tischer, Wilhelm
;  APPLICANT:  Wozny, Manfred
;  TITLE OF INVENTION: Erythropoietin Conjugates
;  FILE REFERENCE: Case 20805
;  CURRENT APPLICATION NUMBER: US/10/014,363
;  CURRENT FILING DATE:  2001-12-11
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 174
;   TYPE: PRT
;   ORGANISM: CHO/dhfr-
US-10-014-363-5

  Query Match            100.0%;  Score 846;  DB 12;  Length 174;
  Best Local Similarity  100.0%;  Pred. No. 9.3e-84;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 68

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 128

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        129 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 173


RESULT 11
US-10-113-824-2
; Sequence 2, Application US/10113824
; Publication No. US20030050269A1
; GENERAL INFORMATION:
;  APPLICANT: Escary, Jean-Louis
;  TITLE OF INVENTION: NEW POLYNUCLEOTIDES AND POLYPEPTIDES OF THE ERYTHROPOIETIN GENE
;  FILE REFERENCE: 021349/0037
;  CURRENT APPLICATION NUMBER: US/10/113,824
;  CURRENT FILING DATE:  2002-03-29
;  PRIOR APPLICATION NUMBER: FR 0104603
;  PRIOR FILING DATE: 2001-04-04
;  PRIOR APPLICATION NUMBER: US 60/343163
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: US 60/345,440
;  PRIOR FILING DATE: 2002-01-04
;  PRIOR APPLICATION NUMBER: US 60/358,598
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-113-824-2

  Query Match            100.0%;  Score 846;  DB 9;  Length 193;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-83;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 147

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        148 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 192


RESULT 12
US-09-813-775C-4
; Sequence 4, Application US/09813775C
; Publication No. US20030054494A1
; GENERAL INFORMATION:
;  APPLICANT: DeSauvage, Frederick
;  APPLICANT:  Henner, Dennis, J.
;  TITLE OF INVENTION: No. US20030054494A1el chimpanzee erythropoietin
;  TITLE OF INVENTION:  polypeptides and nucleic acids encoding the same
;  FILE REFERENCE: GENENT.057CP2
;  CURRENT APPLICATION NUMBER: US/09/813,775C
;  CURRENT FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER: US 09/307307
;  PRIOR FILING DATE: 1999-05-07
;  PRIOR APPLICATION NUMBER: US 09/552265
;  PRIOR FILING DATE: 2000-04-19
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-813-775C-4

  Query Match            100.0%;  Score 846;  DB 9;  Length 193;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-83;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 147

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        148 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 192


RESULT 13
US-09-932-812-22
; Sequence 22, Application US/09932812
; Publication No. US20030082749A1
; GENERAL INFORMATION:
;  APPLICANT: Sun, Lee-Hwei K
```

```
;  APPLICANT: Sun, Bill N
;  APPLICANT: Sun, Cecily R
;  TITLE OF INVENTION: Fc fusion proteins of human erythropoietin with increased biologi
;  FILE REFERENCE: 02SUN2001
;  CURRENT APPLICATION NUMBER: US/09/932,812
;  CURRENT FILING DATE: 2001-10-30
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
;   LENGTH: 435
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: HuEPO-L-vFc gammal with a 27-amino acid leader peptide (Figure 2C
US-09-932-812-22

  Query Match           100.0%;  Score 846;  DB 9;  Length 435;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-83;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 147

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        148 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 192


RESULT 14
US-09-932-812-18
; Sequence 18, Application US/09932812
; Publication No. US20030082749A1
; GENERAL INFORMATION:
;  APPLICANT: Sun, Lee-Hwei K
;  APPLICANT: Sun, Bill N
;  APPLICANT: Sun, Cecily R
;  TITLE OF INVENTION: Fc fusion proteins of human erythropoietin with increased biologi
;  FILE REFERENCE: 02SUN2001
;  CURRENT APPLICATION NUMBER: US/09/932,812
;  CURRENT FILING DATE: 2001-10-30
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
;   LENGTH: 436
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: HuEPO-L-vFc gamma2 with a 27-amino acid  leader peptide (Figure 2
;   OTHER INFORMATION: A)
US-09-932-812-18

  Query Match           100.0%;  Score 846;  DB 9;  Length 436;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-83;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 147

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        148 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 192


RESULT 15
US-09-932-812-20
; Sequence 20, Application US/09932812
; Publication No. US20030082749A1
; GENERAL INFORMATION:
;  APPLICANT: Sun, Lee-Hwei K
;  APPLICANT: Sun, Bill N
;  APPLICANT: Sun, Cecily R
;  TITLE OF INVENTION: Fc fusion proteins of human erythropoietin with increased biol
;  FILE REFERENCE: 02SUN2001
;  CURRENT APPLICATION NUMBER: US/09/932,812
;  CURRENT FILING DATE: 2001-10-30
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
;   LENGTH: 437
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: HuEPO-L-vFc gamma4 with a 27-amino acid leader peptide (Figure
;   OTHER INFORMATION: )
US-09-932-812-20

  Query Match           100.0%;  Score 846;  DB 9;  Length 437;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-83;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 147

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        148 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 192


Search completed: June 13, 2003, 15:13:52
Job time : 48 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:03:00 ; Search time 310 Seconds
(without alignments)
343.164 Million cell updates/sec

Title: US-09-830-964-1
Perfect score: 846
Sequence: 1 APPRLICDSRVLERYLLEAK.....SNFLRGKLKLYTGEACRTGD 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	846	100.0	165	1	PCT-US99-26238-1	Sequence 1, Appli
2	846	100.0	165	1	PCT-US99-26240-1	Sequence 1, Appli
3	846	100.0	165	1	PCT-US99-26241-1	Sequence 1, Appli
4	846	100.0	165	18	US-09-426-566-1	Sequence 1, Appli
5	846	100.0	165	20	US-09-604-938-1	Sequence 1, Appli
6	846	100.0	165	22	US-09-830-964-1	Sequence 1, Appli
7	846	100.0	165	22	US-09-830-967-1	Sequence 1, Appli
8	846	100.0	165	22	US-09-830-968-1	Sequence 1, Appli
9	846	100.0	165	22	US-09-853-731-1	Sequence 1, Appli
10	846	100.0	165	22	US-09-854-018-1	Sequence 1, Appli
11	846	100.0	165	23	US-09-945-517-1	Sequence 1, Appli
12	846	100.0	165	24	US-10-014-363-1	Sequence 1, Appli
13	846	100.0	165	26	US-10-241-356-1	Sequence 1, Appli
14	846	100.0	166	1	PCT-US98-14497-2	Sequence 2, Appli
15	846	100.0	166	3	US-07-766-964A-3	Sequence 3, Appli
16	846	100.0	166	4	US-08-049-802-37	Sequence 37, Appl
17	846	100.0	166	6	US-08-223-263-7	Sequence 7, Appli
18	846	100.0	166	6	US-08-223-263A-7	Sequence 7, Appli
19	846	100.0	166	6	US-08-232-533-37	Sequence 37, Appl
20	846	100.0	166	6	US-08-249-376-7	Sequence 7, Appli
21	846	100.0	166	6	US-08-249-376A-7	Sequence 7, Appli
22	846	100.0	166	6	US-08-249-376B-7	Sequence 7, Appli
23	846	100.0	166	7	US-08-374-540-7	Sequence 7, Appli
24	846	100.0	166	8	US-08-422-020-7	Sequence 7, Appli
25	846	100.0	166	8	US-08-422-020A-7	Sequence 7, Appli
26	846	100.0	166	8	US-08-422-194-7	Sequence 7, Appli
27	846	100.0	166	8	US-08-422-194A-7	Sequence 7, Appli
28	846	100.0	166	8	US-08-422-548-7	Sequence 7, Appli
29	846	100.0	166	8	US-08-422-727-7	Sequence 7, Appli
30	846	100.0	166	8	US-08-422-727A-7	Sequence 7, Appli
31	846	100.0	166	8	US-08-423-194-7	Sequence 7, Appli
32	846	100.0	166	8	US-08-423-194A-7	Sequence 7, Appli
33	846	100.0	166	8	US-08-425-016-7	Sequence 7, Appli
34	846	100.0	166	8	US-08-425-016A-7	Sequence 7, Appli
35	846	100.0	166	8	US-08-425-016B-7	Sequence 7, Appli
36	846	100.0	166	8	US-08-425-020-7	Sequence 7, Appli
37	846	100.0	166	8	US-08-425-020A-7	Sequence 7, Appli
38	846	100.0	166	8	US-08-425-095-7	Sequence 7, Appli
39	846	100.0	166	8	US-08-425-095A-7	Sequence 7, Appli
40	846	100.0	166	8	US-08-433-098A-7	Sequence 7, Appli
41	846	100.0	166	8	US-08-433-098B-7	Sequence 7, Appli
42	846	100.0	166	8	US-08-433-103-7	Sequence 7, Appli
43	846	100.0	166	8	US-08-433-103A-7	Sequence 7, Appli
44	846	100.0	166	8	US-08-433-766-7	Sequence 7, Appli
45	846	100.0	166	8	US-08-433-767-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
PCT-US99-26238-1
; Sequence 1, Application PC/TUS9926238
; GENERAL INFORMATION:
; APPLICANT: Sterrenbeld Biotechnologie North America, Inc.
; APPLICANT: Carcagno, Carlos Miguel
; APPLICANT: Criscuolo, Marcelo
; APPLICANT: Melo, Carlos
; APPLICANT: Vidal, Juan Alejandro
; TITLE OF INVENTION: Host Cells Expressing Recombinant Human Erythropoietin
; FILE REFERENCE: 1792.002PC02
; CURRENT APPLICATION NUMBER: PCT/US99/26238
; CURRENT FILING DATE: 1999-11-08
; EARLIER APPLICATION NUMBER: AR 99-01-00679
; EARLIER FILING DATE: 1999-02-23
; EARLIER APPLICATION NUMBER: AR 98-01-05609
; EARLIER FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US99-26238-1

Query Match 100.0%; Score 846; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-89;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 2
PCT-US99-26240-1
; Sequence 1, Application PC/TUS9926240
; GENERAL INFORMATION:
;  APPLICANT: Sterrenbeld Biotechnologie North America, Inc.
;  APPLICANT: Carcagno, Carlos Miguel
;  APPLICANT: Criscuolo, Marcelo
;  APPLICANT: Melo, Carlos
;  APPLICANT: Vidal, Juan Alejandro
;  TITLE OF INVENTION: Method for the Massive Culture of Cells Producing Recombinant Hum
;  TITLE OF INVENTION: Erythropoietin
;  FILE REFERENCE: 1792.004PC02
;  CURRENT APPLICATION NUMBER: PCT/US99/26240
;  CURRENT FILING DATE: 1999-11-08
;  EARLIER APPLICATION NUMBER: AR 99-01-00681
;  EARLIER FILING DATE: 1999-02-23
;  EARLIER APPLICATION NUMBER: AR 98-01-05611
;  EARLIER FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US99-26240-1

  Query Match           100.0%;  Score 846;  DB 1;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 3
PCT-US99-26241-1
; Sequence 1, Application PC/TUS9926241
; GENERAL INFORMATION:
;  APPLICANT: Sterrenbeld Biotechnologie North America, Inc.
;  APPLICANT: Carcagno, Carlos Miguel
;  APPLICANT: Criscuolo, Marcelo
;  APPLICANT: Melo, Carlos
;  APPLICANT: Vidal, Juan Alejandro
;  TITLE OF INVENTION: Methods of Purifying Recombinant Human Erythropoietin from Cell
;  TITLE OF INVENTION: Culture Supernatants
;  FILE REFERENCE: 1792.003PC02
;  CURRENT APPLICATION NUMBER: PCT/US99/26241
;  CURRENT FILING DATE: 1999-11-08
;  EARLIER APPLICATION NUMBER: AR 99-01-00680
;  EARLIER FILING DATE: 1999-02-23
;  EARLIER APPLICATION NUMBER: AR 98-01-05610
;  EARLIER FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US99-26241-1

  Query Match           100.0%;  Score 846;  DB 1;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 4
US-09-426-566-1
; Sequence 1, Application US/09426566
; GENERAL INFORMATION:
;  APPLICANT: Burke, Paul
;  APPLICANT: Klumb, Lisa
;  APPLICANT: Murphy, Keith
;  APPLICANT: Herberger, John
;  APPLICANT: French, Donna
;  TITLE OF INVENTION: BIODEGRADABLE MICROPARTICLES FOR THE SUSTAINED DELIVERY
;  TITLE OF INVENTION: OF NOVEL ERYTHROPOIETIN STIMULATING PROTEIN
;  FILE REFERENCE: A-626
;  CURRENT APPLICATION NUMBER: US/09/426,566
;  CURRENT FILING DATE: 1999-10-22
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: HUMAN
US-09-426-566-1

  Query Match           100.0%;  Score 846;  DB 18;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 5
US-09-604-938-1
; Sequence 1, Application US/09604938
; GENERAL INFORMATION:
;  APPLICANT:  Bailon, Pascal
;  TITLE OF INVENTION:  ERYTHROPOIETIN CONJUGATES
;  FILE REFERENCE:  1097 nonprovisional
;  CURRENT APPLICATION NUMBER: US/09/604,938
```

```
;  CURRENT FILING DATE:  2000-06-27
;  PRIOR APPLICATION NUMBER:  60/166,151
;  PRIOR FILING DATE:  1999-11-17
;  PRIOR APPLICATION NUMBER:  60/151,548
;  PRIOR FILING DATE:  1999-08-13
;  PRIOR APPLICATION NUMBER:  60/150,225
;  PRIOR FILING DATE:  1999-08-23
;  PRIOR APPLICATION NUMBER:  60/142,254
;  PRIOR FILING DATE:  1999-07-02
;  NUMBER OF SEQ ID NOS:  3
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-604-938-1

  Query Match           100.0%;  Score 846;  DB 20;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 6
US-09-830-964-1
; Sequence 1, Application US/09830964
; GENERAL INFORMATION:
;  APPLICANT: Sterrenbeld Biotechnologie North America, Inc.
;  APPLICANT:  Carcagno, Carlos Miguel
;  APPLICANT:  Criscuolo, Marcelo
;  APPLICANT:  Melo, Carlos
;  APPLICANT:  Vidal, Juan Alejandro
;  TITLE OF INVENTION: Methods of Purifying Recombinant Human
;  TITLE OF INVENTION:  Erythropoietin from Cell Culture Supernatants
;  FILE REFERENCE: 1909.0030002
;  CURRENT APPLICATION NUMBER: US/09/830,964
;  CURRENT FILING DATE:  1999-11-08
;  PRIOR APPLICATION NUMBER: AR 99-01-00680
;  PRIOR FILING DATE: 1999-02-23
;  PRIOR APPLICATION NUMBER: AR 98-01-05610
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-830-964-1

  Query Match           100.0%;  Score 846;  DB 22;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 7
US-09-830-967-1
; Sequence 1, Application US/09830967
; GENERAL INFORMATION:
;  APPLICANT: Sterrenbeld Biotechnologie North America, Inc.
;  APPLICANT:  Carcagno, Carlos Miguel
;  APPLICANT:  Criscuolo, Marcelo
;  APPLICANT:  Melo, Carlos
;  APPLICANT:  Vidal, Juan Alejandro
;  TITLE OF INVENTION: Host Cells Expressing Recombinant Human Erythropoietin
;  FILE REFERENCE: 1909.0020002
;  CURRENT APPLICATION NUMBER: US/09/830,967
;  CURRENT FILING DATE:  1999-11-08
;  PRIOR APPLICATION NUMBER: AR 99-01-00679
;  PRIOR FILING DATE: 1999-02-23
;  PRIOR APPLICATION NUMBER: AR 98-01-05609
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-830-967-1

  Query Match           100.0%;  Score 846;  DB 22;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 8
US-09-830-968-1
; Sequence 1, Application US/09830968
; GENERAL INFORMATION:
;  APPLICANT: Sterrenbeld Biotechnologie North America, Inc.
;  APPLICANT:  Carcagno, Carlos Miguel
;  APPLICANT:  Criscuolo, Marcelo
;  APPLICANT:  Melo, Carlos
;  APPLICANT:  Vidal, Juan Alejandro
;  TITLE OF INVENTION: Method for the Massive Culture of Cells
;  TITLE OF INVENTION:  Producing Recombinant Human Erythropoietin
;  FILE REFERENCE: 1909.0040002
;  CURRENT APPLICATION NUMBER: US/09/830,968
;  CURRENT FILING DATE:  1999-11-08
;  PRIOR APPLICATION NUMBER: AR 99-01-00681
;  PRIOR FILING DATE: 1999-02-23
;  PRIOR APPLICATION NUMBER: AR 98-01-05611
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-830-968-1

  Query Match           100.0%;  Score 846;  DB 22;  Length 165;
```

```
  Best Local Similarity   100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 9
US-09-853-731-1
; Sequence 1, Application US/09853731
; GENERAL INFORMATION:
;  APPLICANT: Papadimitriou, Apollon
;  TITLE OF INVENTION: Erythropoietin Composition
;  FILE REFERENCE: 20619 US
;  CURRENT APPLICATION NUMBER: US/09/853,731
;  CURRENT FILING DATE:  2001-05-11
;  PRIOR APPLICATION NUMBER: EP/00110355.5
;  PRIOR FILING DATE: 2000-05-15
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-853-731-1

  Query Match            100.0%;  Score 846;  DB 22;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 10
US-09-854-018-1
; Sequence 1, Application US/09854018
; GENERAL INFORMATION:
;  APPLICANT: Herberger, John
;  TITLE OF INVENTION: Secondary Drying of Microparticles Using Pressurized Carbon Dioxi
;  FILE REFERENCE: A-786
;  CURRENT APPLICATION NUMBER: US/09/854,018
;  CURRENT FILING DATE:  2001-05-10
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Human
US-09-854-018-1

  Query Match            100.0%;  Score 846;  DB 22;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 11
US-09-945-517-1
; Sequence 1, Application US/09945517
; GENERAL INFORMATION:
;  APPLICANT: Li, Tiansheng
;  APPLICANT:  Chang, Byeong
;  APPLICANT:  Sloey, Christopher
;  TITLE OF INVENTION: L-METHIONINE AS A STABILIZER FOR NESP/EPO IN HSA-FREE FORMULAT
;  FILE REFERENCE: A-803
;  CURRENT APPLICATION NUMBER: US/09/945,517
;  CURRENT FILING DATE:  2001-08-30
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-945-517-1

  Query Match            100.0%;  Score 846;  DB 23;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 12
US-10-014-363-1
; Sequence 1, Application US/10014363
; GENERAL INFORMATION:
;  APPLICANT: Burg, Josef
;  APPLICANT:  Engel, Alfred
;  APPLICANT:  Franze, Reinhard
;  APPLICANT:  Hilger, Bernd
;  APPLICANT:  Schurig, Hartmut Ernst
;  APPLICANT:  Tischer, Wilhelm
;  APPLICANT:  Wozny, Manfred
;  TITLE OF INVENTION: Erythropoietin Conjugates
;  FILE REFERENCE: Case 20805
;  CURRENT APPLICATION NUMBER: US/10/014,363
;  CURRENT FILING DATE:  2001-12-11
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-014-363-1

  Query Match            100.0%;  Score 846;  DB 24;  Length 165;
```

```
  Best Local Similarity   100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 13
US-10-241-356-1
; Sequence 1, Application US/10241356
; GENERAL INFORMATION:
;  APPLICANT: TISCHER, WILHELM
;  TITLE OF INVENTION: DIGLYCOSYLATED ERYTHROPOIETIN
;  FILE REFERENCE: 20971
;  CURRENT APPLICATION NUMBER: US/10/241,356
;  CURRENT FILING DATE:  2002-09-11
;  PRIOR APPLICATION NUMBER: EP 01122555.4
;  PRIOR FILING DATE: 2001-09-25
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-241-356-1

  Query Match             100.0%;  Score 846;  DB 26;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 14
PCT-US98-14497-2
; Sequence 2, Application PC/TUS9814497
; GENERAL INFORMATION:
;  APPLICANT: Cox III, George N
;  APPLICANT: Bolder Biotechnology, Inc.
;  TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
;  FILE REFERENCE: BB0011
;  CURRENT APPLICATION NUMBER: PCT/US98/14497
;  CURRENT FILING DATE: 1998-07-13
;  EARLIER APPLICATION NUMBER: 60/052,516
;  EARLIER FILING DATE: 1997-07-14
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US98-14497-2

  Query Match             100.0%;  Score 846;  DB 1;  Length 166;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 15
US-07-766-964A-3
; Sequence 3, Application US/07766964A
;  GENERAL INFORMATION:
;    APPLICANT:  Rosen, Jonathan I.
;    TITLE OF INVENTION:  HYBRID GROWTH FACTORS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Robert L. Minier
;      STREET:  1 Johnson & Johnson Plaza
;      CITY:  New Brunswick
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  08933
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/766,964A
;      FILING DATE:  19910926
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/589,958
;      FILING DATE:  28-SEP-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stark, Michael
;      REGISTRATION NUMBER:  32,495
;      REFERENCE/DOCKET NUMBER:  BCI-15
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-524-2817
;      TELEFAX:  908-524-2808
;      TELEX:  844-481
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-07-766-964A-3

  Query Match             100.0%;  Score 846;  DB 3;  Length 166;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-89;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
```

```
Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
```

Search completed: June 13, 2003, 15:10:21
Job time : 311 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:03:51 ; Search time 108 Seconds
(without alignments)
389.731 Million cell updates/sec

Title: US-09-830-964-1
Perfect score: 846
Sequence: 1 APPRLICDSRVLERYLLEAK.....SNFLRGKLKLYTGEACRTGD 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1155225 seqs, 255096386 residues

Total number of hits satisfying chosen parameters: 1155225

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	846	100.0	165	5	US-09-687-981-1	Sequence 1, Appli
2	846	100.0	165	6	US-10-293-551-1	Sequence 1, Appli
3	846	100.0	165	6	US-10-411-012-73	Sequence 73, Appl
4	846	100.0	165	6	US-10-411-026-73	Sequence 73, Appl
5	846	100.0	165	6	US-10-410-930-73	Sequence 73, Appl
6	846	100.0	165	6	US-10-410-945-73	Sequence 73, Appl
7	846	100.0	165	6	US-10-410-962-73	Sequence 73, Appl
8	846	100.0	165	6	US-10-410-980-73	Sequence 73, Appl
9	846	100.0	165	6	US-10-411-044-73	Sequence 73, Appl
10	846	100.0	165	6	US-10-411-049-73	Sequence 73, Appl
11	846	100.0	165	6	US-10-410-913-73	Sequence 73, Appl
12	846	100.0	165	6	US-10-411-037-73	Sequence 73, Appl
13	846	100.0	165	6	US-10-411-043-73	Sequence 73, Appl
14	846	100.0	165	6	US-10-410-897-73	Sequence 73, Appl
15	846	100.0	165	6	US-10-410-997-73	Sequence 73, Appl
16	846	100.0	166	5	US-09-708-506A-2	Sequence 2, Appli
17	846	100.0	166	6	US-10-293-551-2	Sequence 2, Appli
18	846	100.0	166	6	US-10-298-148-2	Sequence 2, Appli
19	846	100.0	166	6	US-10-400-377-2	Sequence 2, Appli
20	846	100.0	166	6	US-10-400-708-2	Sequence 2, Appli
21	846	100.0	166	6	US-10-360-101-227	Sequence 227, App
22	846	100.0	192	1	PCT-US02-40891-593	Sequence 593, App
23	846	100.0	192	1	PCT-US02-40891-594	Sequence 594, App
24	846	100.0	192	1	PCT-US02-40891-603	Sequence 603, App
25	846	100.0	192	1	PCT-US02-40891-1689	Sequence 1689, Ap
26	846	100.0	192	1	PCT-US02-40891-1690	Sequence 1690, Ap
27	846	100.0	192	1	PCT-US02-40891-1691	Sequence 1691, Ap
28	846	100.0	192	1	PCT-US02-40891-1828	Sequence 1828, Ap
29	846	100.0	192	1	PCT-US02-40891-1829	Sequence 1829, Ap
30	846	100.0	192	1	PCT-US02-40891-1830	Sequence 1830, Ap
31	846	100.0	193	5	US-09-728-403B-1	Sequence 1, Appli
32	846	100.0	193	5	US-09-723-955-1	Sequence 1, Appli
33	846	100.0	193	7	US-60-455-444-5994	Sequence 5994, Ap
34	846	100.0	193	7	US-60-465-241-5994	Sequence 5994, Ap
35	846	100.0	209	6	US-10-230-454-4	Sequence 4, Appli
36	846	100.0	370	6	US-10-230-454-3	Sequence 3, Appli
37	846	100.0	428	1	PCT-US03-14428-10	Sequence 10, Appl
38	846	100.0	428	6	US-10-435-608-10	Sequence 10, Appl
39	846	100.0	435	5	US-09-932-812A-22	Sequence 22, Appl
40	846	100.0	436	5	US-09-932-812A-18	Sequence 18, Appl
41	846	100.0	437	5	US-09-932-812A-20	Sequence 20, Appl
42	846	100.0	768	1	PCT-US02-40891-1521	Sequence 1521, Ap
43	846	100.0	768	1	PCT-US02-40891-1522	Sequence 1522, Ap
44	846	100.0	768	1	PCT-US02-40891-1523	Sequence 1523, Ap
45	846	100.0	768	1	PCT-US02-40891-1660	Sequence 1660, Ap

ALIGNMENTS

RESULT 1
US-09-687-981-1
; Sequence 1, Application US/09687981
; GENERAL INFORMATION:
; APPLICANT: Burke, Paul
; APPLICANT: Klumb, Lisa
; APPLICANT: Murphy, Keith
; APPLICANT: Herberger, John
; APPLICANT: French, Donna
; TITLE OF INVENTION: BIODEGRADABLE MICROPARTICLES FOR THE SUSTAINED DELIVERY OF NOV
; TITLE OF INVENTION: ERYTHROPOIETIN STIMULATING PROTEIN
; FILE REFERENCE: A-626A
; CURRENT APPLICATION NUMBER: US/09/687,981
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/426,566
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-687-981-1

Query Match 100.0%; Score 846; DB 5; Length 165;
Best Local Similarity 100.0%; Pred. No. 7.2e-78;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
```

RESULT 2
US-10-293-551-1
; Sequence 1, Application US/10293551
; GENERAL INFORMATION:
; APPLICANT: Bailon, Pascal
; TITLE OF INVENTION: ERYTHROPOIETIN CONJUGATES
; FILE REFERENCE: 1097 nonprovisional
; CURRENT APPLICATION NUMBER: US/10/293,551

```
;  CURRENT FILING DATE:  2002-11-14
;  PRIOR APPLICATION NUMBER: US/09/604,938
;  PRIOR FILING DATE: 2000-06-27
;  PRIOR APPLICATION NUMBER: 60/166,151
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: 60/151,548
;  PRIOR FILING DATE: 1999-08-13
;  PRIOR APPLICATION NUMBER: 60/150,225
;  PRIOR FILING DATE: 1999-08-23
;  PRIOR APPLICATION NUMBER: 60/142,254
;  PRIOR FILING DATE: 1999-07-02
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-293-551-1

  Query Match             100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 3
US-10-411-012-73
; Sequence 73, Application US/10411012
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryne
;  TITLE OF INVENTION: GLYCOPEGYLATION METHODS AND PROTEINS/PEPTIDES PRODUCED BY THE
;  TITLE OF INVENTION:  METHODS
;  FILE REFERENCE: 040853-01-5051
;  CURRENT APPLICATION NUMBER: US/10/411,012
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-411-012-73

  Query Match             100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 4
US-10-411-026-73
; Sequence 73, Application US/10411026
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  TITLE OF INVENTION: PROTEIN REMODELING METHODS AND PROTEINS/PEPTIDES PRODUCED BY T
;  TITLE OF INVENTION:  METHODS
;  FILE REFERENCE: 040853-01-5053
;  CURRENT APPLICATION NUMBER: US/10/411,026
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-411-026-73

  Query Match             100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 5
```

```
US-10-410-930-73
; Sequence 73, Application US/10410930
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryn
;  TITLE OF INVENTION: INTERFERON BETA: REMODELING AND GLYCOCONJUGATION OF INTERFERON
;  TITLE OF INVENTION:  BETA
;  FILE REFERENCE: 040853-01-5056
;  CURRENT APPLICATION NUMBER: US/10/410,930
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-410-930-73

  Query Match             100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 6
US-10-410-945-73
; Sequence 73, Application US/10410945
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryn
;  TITLE OF INVENTION: ERYTHROPOIETIN: REMODELING AND GLYCOCONJUGATION OF ERYTHROPOIETIN
;  FILE REFERENCE: 040853-01-5083
;  CURRENT APPLICATION NUMBER: US/10/410,945
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-410-945-73

  Query Match             100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 7
US-10-410-962-73
; Sequence 73, Application US/10410962
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryn
;  TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR: REMODELING AND
;  TITLE OF INVENTION:  GLYCOCONJUGATION OF G-CSF
;  FILE REFERENCE: 040853-01-5054
;  CURRENT APPLICATION NUMBER: US/10/410,962
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-410-962-73
```

```
  Query Match          100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 8
US-10-410-980-73
; Sequence 73, Application US/10410980
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryn
;  TITLE OF INVENTION: INTERLEUKIN-2:  REMODELING AND GLYCOCONJUGATION OF IL-2
;  FILE REFERENCE: 040853-01-5066
;  CURRENT APPLICATION NUMBER: US/10/410,980
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-410-980-73

  Query Match          100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 9
US-10-411-044-73
; Sequence 73, Application US/10411044
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryn
;  TITLE OF INVENTION: BETA-GLUCOSIDASE: REMODELING AND GLYCOCONJUGATION OF
;  TITLE OF INVENTION:  BETA-GLUCOSIDASE
;  FILE REFERENCE: 040853-01-5064
;  CURRENT APPLICATION NUMBER: US/10/411,044
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-411-044-73

  Query Match          100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 10
US-10-411-049-73
; Sequence 73, Application US/10411049
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryn
;  TITLE OF INVENTION: INTERFERON ALPHA: REMODELING AND GLYCOCONJUGATION OF INTERFERO
;  TITLE OF INVENTION:  ALPHA
;  FILE REFERENCE: 040853-01-5055
;  CURRENT APPLICATION NUMBER: US/10/411,049
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
```

```
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-411-049-73

  Query Match             100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 11
US-10-410-913-73
; Sequence 73, Application US/10410913
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryn
;  TITLE OF INVENTION: GLYCOCONJUGATION METHODS AND PROTEINS/PEPTIDES PRODUCED BY THE
;  TITLE OF INVENTION:  METHODS
;  FILE REFERENCE: 040853-01-5081
;  CURRENT APPLICATION NUMBER: US/10/410,913
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-410-913-73

  Query Match             100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 12
US-10-411-037-73
; Sequence 73, Application US/10411037
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryn
;  TITLE OF INVENTION: ALPHA GALACTOSIDASE A: REMODELING AND GLYCOCONJUGATION OF ALPH
;  TITLE OF INVENTION:  GALACTOSIDASE A
;  FILE REFERENCE: 040853-01-5082
;  CURRENT APPLICATION NUMBER: US/10/411,037
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-411-037-73

  Query Match             100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 13
```

```
US-10-411-043-73
; Sequence 73, Application US/10411043
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryn
;  TITLE OF INVENTION: ANTI-TUMOR NECROSIS ALPHA: REMODELING AND GLYCOCONJUGATION OF
;  TITLE OF INVENTION:  ANTI-TUMOR NECROSIS FACTOR ALPHA
;  FILE REFERENCE: 040853-01-5075
;  CURRENT APPLICATION NUMBER: US/10/411,043
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-411-043-73

  Query Match             100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 14
US-10-410-897-73
; Sequence 73, Application US/10410897
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryn
;  TITLE OF INVENTION: FACTOR IX; REMODELING AND GLYCOCONJUGATION OF FACTOR IX
;  FILE REFERENCE: 040853-01-5058
;  CURRENT APPLICATION NUMBER: US/10/410,897
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-410-897-73

  Query Match             100.0%;  Score 846;  DB 6;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-78;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165


RESULT 15
US-10-410-997-73
; Sequence 73, Application US/10410997
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  APPLICANT:  Bowe, Caryn
;  TITLE OF INVENTION: FOLLICLE STIMULATING HORMONE: REMODELING AND GLYCOCONJUGATION
;  TITLE OF INVENTION:  FSH
;  FILE REFERENCE: 040853-01-5059
;  CURRENT APPLICATION NUMBER: US/10/410,997
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-410-997-73
```

Query Match 100.0%; Score 846; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 7.2e-78;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
```

Search completed: June 13, 2003, 15:12:17
Job time : 109 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:03:00 ; Search time 24 Seconds
(without alignments)
660.924 Million cell updates/sec

Title: US-09-830-964-1
Perfect score: 846
Sequence: 1 APPRLICDSRVLERYLLEAK.....SNFLRGKLKLYTGEACRTGD 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	846	100.0	193	1	ZUHU	erythropoietin pre
2	764.5	90.4	192	1	JQ0173	erythropoietin pre
3	759.5	89.8	192	1	I84613	erythropoietin pre
4	713	84.3	188	1	I46083	erythropoietin pre
5	701	82.9	192	1	S28148	erythropoietin pre
6	685.5	81.0	194	1	I46401	erythropoietin pre
7	681	80.5	192	1	A24902	erythropoietin pre
8	680.5	80.4	195	2	JC7699	erythropoietin - r
9	678	80.1	190	2	I46578	erythropoietin - p
10	638	75.4	175	2	I46199	erythropoietin - d
11	90	10.6	353	2	G02729	thrombopoietin - h
12	89	10.5	353	2	I80105	thrombopoietin pre
13	88	10.4	323	2	AB0323	ribonucleoside-dip
14	87.5	10.3	346	2	AE0959	Solute binding rec
15	86	10.2	286	2	A55530	megakaryocyte grow
16	83	9.8	296	2	AI0443	probable 2-hydroxy
17	83	9.8	339	2	A83274	UDP-N-acetylpyruvo
18	80.5	9.5	3033	1	GNWVJ8	genome polyprotein
19	79.5	9.4	1829	2	T35681	probable sensory h
20	79	9.3	480	2	S56639	ribosomal protein
21	78.5	9.3	813	2	AF0526	ATP-dependent heli
22	78.5	9.3	897	2	A54696	EGF receptor subst
23	78	9.2	348	2	T35450	ABC transporter AT
24	78	9.2	455	2	H97693	methylamine utiliz
25	78	9.2	455	2	AG2919	conserved hypothet
26	78	9.2	747	1	S36741	probable copper-tr
27	77.5	9.2	242	2	AD1928	hypothetical prote
28	77	9.1	451	2	S75569	hypothetical prote
29	76.5	9.0	154	2	H82810	bacterioferritin X
30	76.5	9.0	425	2	AE3465	mandelate racemase
31	75.5	8.9	637	2	S75772	hypothetical prote
32	74.5	8.8	400	2	AB2922	conserved hypothet
33	74.5	8.8	425	2	C97696	rtS beta (AF305057
34	74.5	8.8	824	2	D64738	ATP-dependent heli
35	74	8.7	282	2	B37994	RF2 protein - saim
36	74	8.7	326	2	JC4125	thrombopoietin pre
37	74	8.7	335	2	AH3625	ribonucleoside-dip
38	74	8.7	1564	2	S55517	probable transport
39	73.5	8.7	401	2	H83911	hypothetical prote
40	73.5	8.7	476	1	S71789	GCN5 protein - hum
41	73.5	8.7	717	2	F82613	VacB protein XF198
42	73	8.6	263	2	B75361	WD-repeat family p
43	73	8.6	1089	2	S53978	PSE1 protein - yea
44	72.5	8.6	379	2	H69478	NADH2 dehydrogenas
45	72.5	8.6	401	2	AF3341	precorrin-6y c5,15

ALIGNMENTS

RESULT 1
ZUHU
erythropoietin precursor [validated] - human
C;Species: Homo sapiens (man)
C;Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000
C;Accession: A01855; A24744; A25384; A22210; S56178
R;Jacobs, K.; Shoemaker, C.; Rudersdorf, R.; Neill, S.D.; Kaufman, R.J.; Mufson, A.;
Nature 313, 806-810, 1985
A;Title: Isolation and characterization of genomic and cDNA clones of human erythropo
A;Reference number: A01855; MUID:85137899; PMID:3838366
A;Accession: A01855
A;Molecule type: mRNA; DNA
A;Residues: 1-193 <JAC>
A;Cross-references: GB:X02157; GB:X02158
R;Lin, F.K.; Suggs, S.; Lin, C.H.; Browne, J.K.; Smalling, R.; Egrie, J.C.; Chen, K.K
Proc. Natl. Acad. Sci. U.S.A. 82, 7580-7584, 1985
A;Title: Cloning and expression of the human erythropoietin gene.
A;Reference number: A24744; MUID:86067948; PMID:3865178
A;Accession: A24744
A;Molecule type: DNA
A;Residues: 1-193 <LIN>
A;Cross-references: GB:M11319; NID:g182197; PIDN:AAA52400.1; PID:g182198
R;Lai, P.H.; Everett, R.; Wang, F.F.; Arakawa, T.; Goldwasser, E.
J. Biol. Chem. 261, 3116-3121, 1986
A;Title: Structural characterization of human erythropoietin.
A;Reference number: A25384; MUID:86140080; PMID:3949763
A;Accession: A25384
A;Molecule type: protein
A;Residues: 28-86,'Q',87-193 <LAI>
A;Experimental source: urine
A;Note: forms without the carboxyl-terminal residue and the four carboxyl-terminal re
R;Yanagawa, S.; Hirade, K.; Ohnota, H.; Sasaki, R.; Chiba, H.; Ueda, M.; Goto, M.
J. Biol. Chem. 259, 2707-2710, 1984
A;Title: Isolation of human erythropoietin with monoclonal antibodies.
A;Reference number: A22210; MUID:84135751; PMID:6698989
A;Accession: A22210
A;Molecule type: protein
A;Residues: 28-29,'X',31-33,'L',35-50,'X',52-53,'D',55,'G',57 <YAN>
R;Matsumoto, S.; Ikura, K.; Ueda, M.; Sasaki, R.
Plant Mol. Biol. 27, 1163-1172, 1995
A;Title: Characterization of a human glycoprotein (erythropoietin) produced in cultur
A;Reference number: S56178; MUID:95284365; PMID:7766897
A;Accession: S56178
A;Molecule type: protein
A;Residues: 28-33,'X',35-37 <MTS>
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C;Genetics:
A;Gene: GDB:EPO
A;Cross-references: GDB:119110; OMIM:133170
A;Map position: 7q21.3-7q22.1
A;Introns: 5/1; 53/3; 82/3; 142/3
C;Function:

```
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-193/Product: erythropoietin #status experimental <MAT>
F;34-188,56-60/Disulfide bonds: #status experimental
F;51,65,110/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;153/Binding site: carbohydrate (Ser) (covalent) #status experimental

  Query Match             100.0%;  Score 846;  DB 1;  Length 193;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-74;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 147

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        148 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 192


RESULT 2
JQ0173
erythropoietin precursor - crab-eating macaque
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 07-Sep-1990 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C;Accession: JQ0173
R;Lin, F.K.; Lin, C.H.; Lai, P.H.; Browne, J.K.; Egrie, J.C.; Smalling, R.; Fox, G.M.; C
Gene 44, 201-209, 1986
A;Title: Monkey erythropoietin gene: cloning, expression and comparison with the human e
A;Reference number: JQ0173; MUID:87055236; PMID:2877922
A;Accession: JQ0173
A;Molecule type: mRNA
A;Residues: 1-192 <LIN>
A;Cross-references: GB:M18189; GB:M15818; GB:M15819; GB:M18188; NID:g342093; PIDN:AAA368
A;Experimental source: kidney
C;Comment: This protein is the principal hormone involved in the regulation of erythrocy
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by liver c
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-192/Product: erythropoietin #status predicted <MAT>
F;34-187,56-60/Disulfide bonds: #status predicted
F;51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;152/Binding site: carbohydrate (Ser) (covalent) #status predicted

  Query Match             90.4%;  Score 764.5;  DB 1;  Length 192;
  Best Local Similarity   91.5%;  Pred. No. 1.4e-66;
  Matches  151;  Conservative    7;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||||||||||||||||||||:| ||:| ||||||||||||||||||||||||||||
Db         28 APPRLICDSRVLERYLLEAKEAENVTMGCSESCSLNENITVPDTKVNFYAWKRMEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||:| ||||:||||||:|||:|||||:|||||||||| ||||
Db         88 VEVWQGLALLSEAVLRGQAVLANSSQPFEPLQLHMDKAISGLRSITTLLRALGAQ-EAIS 146

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||| |||||||||||||||||| ||
Db        147 LPDAASAAPLRTITADTFCKLFRVYSNFLRGKLKLYTGEACRRGD 191


RESULT 3
I84613
erythropoietin precursor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 02-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C;Accession: I84613
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence hom
A;Reference number: I46083; MUID:93372347; PMID:8364201
A;Accession: I84613
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-192 <RES>
A;Cross-references: GB:L10609; NID:g342095; PIDN:AAA36842.1; PID:g342096
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-192/Product: erythropoietin #status predicted <MAT>
F;34-187,56-60/Disulfide bonds: #status predicted
F;51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;152/Binding site: carbohydrate (Ser) (covalent) #status predicted

  Query Match             89.8%;  Score 759.5;  DB 1;  Length 192;
  Best Local Similarity   90.3%;  Pred. No. 4.3e-66;
  Matches  149;  Conservative    9;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||:|||||||||||||||||:| ||:| ||||||||||||||||||||:||||||
Db         28 APPRLVCDSRVLERYLLEAKEAENVTMGCSESCSLNENITVPDTKVNFYAWKRIEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||:| |||||:||||||:|||:|||||:|||||||||| ||||
Db         88 VEVWQGLALLSEAVLRGQAVLANSSQPFEPLQLHMDKAISGLRSITTLLRALGAQ-EAIS 146

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||| |||||||||||||||||| ||
Db        147 LPDAASAAPLRTITADTFCKLFRVYSNFLRGKLKLYTGEACRRGD 191


RESULT 4
I46083
erythropoietin precursor - cat (fragment)
C;Species: Felis silvestris catus (domestic cat)
C;Date: 16-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C;Accession: I46083
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence hom
A;Reference number: I46083; MUID:93372347; PMID:8364201
A;Accession: I46083
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-188 <WEN>
A;Cross-references: GB:L10606; NID:g163820; PIDN:AAA30807.1; PID:g163821
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-22/Domain: signal sequence (fragment) #status predicted <SIG>
F;23-188/Product: erythropoietin #status predicted <MAT>
F;29-183,51-55/Disulfide bonds: #status predicted
F;46,60,105/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;148/Binding site: carbohydrate (Ser) (covalent) #status predicted

  Query Match             84.3%;  Score 713;  DB 1;  Length 188;
  Best Local Similarity   84.2%;  Pred. No. 1.3e-61;
  Matches  139;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||:|||:||||:| |||| || :|||||||||||||| ||||:|||||
Db         23 APPRLICDSRVLERYILEAREAENVTMGCAEGCSFSENITVPDTKVNFYTWKRMDVGQQA 82
```

```
Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||:|||||||  |||||  |  |||||||||  |||||:||||||||||  |
Db         83 VEVWQGLALLSEAILRGQALLANSSQPSETLQLHVDKAVSSLRSLTSLLRALGAQKEATS 142

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
               |:|  |||||||  |  ||  ||||:|||||||||  ||||||||  ||
Db        143 LPEATSAAPLRTFTVDTLCKLFRIYSNFLRGKLTLYTGEACRRGD 187


RESULT 5
S28148
erythropoietin precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 15-Nov-1996 #text_change 21-Jul-2000
C;Accession: S28148; I62743
R;Nagao, M.; Suga, H.; Okano, M.; Masuda, S.; Narita, H.; Ikura, K.; Sasaki, R.
Biochim. Biophys. Acta 1171, 99-102, 1992
A;Title: Nucleotide sequence of rat erythropoietin.
A;Reference number: S28148; MUID:93042015; PMID:1420369
A;Accession: S28148
A;Molecule type: mRNA
A;Residues: 1-192 <NAG>
A;Cross-references: GB:D10763; NID:g220735; PIDN:BAA01593.1; PID:g220736
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence homolo
A;Reference number: I46083; MUID:93372347; PMID:8364201
A;Accession: I62743
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 4-192 <RES>
A;Cross-references: GB:L10608; NID:g204060; PIDN:AAA41126.1; PID:g204061
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by liver o
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-192/Product: erythropoietin #status predicted <MAT>
F;33-187,55-165/Disulfide bonds: #status predicted
F;50,64,109/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             82.9%;  Score 701;  DB 1;  Length 192;
  Best Local Similarity   82.4%;  Pred. No. 2e-60;
  Matches  136;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||:||||||||:| ||||   |:||||||||||||||||||:| :||
Db         27 APPRLICDSRVLERYILEAKEAENVTMGCAEGPRLSENITVPDTKVNFYAWKRMKVEEQA 86

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              |||||||:|||||:|: |||  ||||| | ||||:|||:|||||||:||| ||||||  :|
Db         87 VEVWQGLSLLSEAILQAQALQANSSQPPESLQLHIDKAISGLRSLTSLLRVLGAQKELMS 146

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              ||||  ||||||:||||| |||||||||||||||||||||||| ||
Db        147 PPDATQAAPLRTLTADTFCKLFRVYSNFLRGKLKLYTGEACRRGD 191


RESULT 6
I46401
erythropoietin precursor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 16-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C;Accession: I46401; I47077
R;Fu, P.; Evans, B.; Lim, G.B.; Moritz, K.; Wintour, E.M.
Mol. Cell. Endocrinol. 93, 107-116, 1993
A;Title: The sheep erythropoietin gene: molecular cloning and effect of hemorrhage on pl
A;Reference number: I46401; MUID:93351736; PMID:8349021
A;Accession: I46401
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-194 <FUX>
A;Cross-references: EMBL:Z24681; NID:g395049; PIDN:CAA80848.1; PID:g395050
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence hom
A;Reference number: I46083; MUID:93372347; PMID:8364201
A;Accession: I47077
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 4-15,'L',17-107,'P',109-194 <WEN>
A;Cross-references: GB:L10610; NID:g165876; PIDN:AAA31518.1; PID:g165877
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-194/Product: erythropoietin #status predicted <MAT>
F;34-189,56-60/Disulfide bonds: #status predicted
F;51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;154/Binding site: carbohydrate (Ser) (covalent) #status predicted

  Query Match             81.0%;  Score 685.5;  DB 1;  Length 194;
  Best Local Similarity   81.9%;  Pred. No. 6.3e-59;
  Matches  136;  Conservative    9;  Mismatches   20;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||:|||:|||| | |||| || :|||||||||||||||||||| |||
Db         28 APPRLICDSRVLERYILEAREAENATMGCAEGCSFSENITVPDTKVNFYAWKRMEVQQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :||||||||||||: |||||| |:||| | |:|||||||||||||:||||||||||||
Db         88 LEVWQGLALLSEAIFRGQALLANASQPCEALRLHVDKAVSGLRSLTSLLRALGAQKEAIP 147

Qy        121 PPDAA-SAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
               |||  ||||||  | |   ||||:||||||||| ||||||||  ||
Db        148 LPDATPSAAPLRIFTVDALSKLFRIYSNFLRGKLTLYTGEACRRGD 193


RESULT 7
A24902
erythropoietin precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 25-Oct-1987 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C;Accession: A24902; A24901
R;Shoemaker, C.B.; Mitsock, L.D.
Mol. Cell. Biol. 6, 849-858, 1986
A;Title: Murine erythropoietin gene: cloning, expression, and human gene homology.
A;Reference number: A24902; MUID:87039105; PMID:3773894
A;Accession: A24902
A;Molecule type: DNA
A;Residues: 1-192 <SHO>
A;Note: the authors translated the codon TTA for residue 12 as Phe, TTA for residue 4
R;McDonald, J.D.; Lin, F.K.; Goldwasser, E.
Mol. Cell. Biol. 6, 842-848, 1986
A;Title: Cloning, sequencing, and evolutionary analysis of the mouse erythropoietin g
A;Reference number: A24901; MUID:87039104; PMID:3022133
A;Accession: A24901
A;Molecule type: DNA
A;Residues: 1-67,'P',69-192 <MCD>
A;Cross-references: GB:M12930; NID:g193086; PIDN:AAA37570.1; PID:g387152
C;Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C;Genetics:
A;Introns: 5/1; 52/3; 81/3; 141/3
C;Function:
A;Description: the primary inducer of erythrocyte formation
C;Superfamily: erythropoietin
C;Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-192/Product: erythropoietin #status predicted <MAT>
F;33-187,55-165/Disulfide bonds: #status predicted
F;50,64,109/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
  Query Match            80.5%;  Score 681;  DB 1;  Length 192;
  Best Local Similarity  79.4%;  Pred. No. 1.7e-58;
  Matches  131;  Conservative   14;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||:||||||||:| ||||   |:|||||| |||||||||||||| :||
Db         27 APPRLICDSRVLERYILEAKEAENVTMGCAEGPRLSENITVADTKVNFYAWKRMEVEEQA 86

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :|||||:|||||:|: |||| ||||| | ||||:|||:|||||||:||| |||||| :|
Db         87 IEVWQGLSLLSEAILQAQALLANSSQPPETLQLHIDKAISGLRSLTSLLRVLGAQKELMS 146

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||    |||||:| ||| |||||||:||||||||||||| || ||
Db        147 PPDTTPPAPLRTLTVDTFCKLFRVYANFLRGKLKLYTGEVCRRGD 191


RESULT 8
JC7699
erythropoietin - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C;Accession: JC7699
R;Vilalta, A.; Wu, D.; Margalith, M.; Hobart, P.
Biochem. Biophys. Res. Commun. 284, 823-827, 2001
A;Title: Rabbit EPO gene and cDNA: Expression of rabbit EPO after intramuscular injectic
A;Reference number: JC7699; MUID:21290682; PMID:11396976
A;Contents: Kidney
A;Accession: JC7699
A;Molecule type: DNA
A;Residues: 1-195 <VIL>
A;Cross-references: GB:AF290943
C;Comment: This protein, a heavily glycosylated 34K protein produced in the fetal liver
cytes.
C;Genetics:
A;Gene: epo
C;Superfamily: erythropoietin
C;Keywords: glycoprotein; kidney

  Query Match            80.4%;  Score 680.5;  DB 2;  Length 195;
  Best Local Similarity  81.3%;  Pred. No. 1.9e-58;
  Matches  135;  Conservative   12;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              || |||||||||||:||||||||:| |||| ||| |||||||||||: ||: | |: |
Db         29 APARLICDSRVLERYILEAKEAENVTMGCAEGCSLGENITVPDTKVNFHHWKKSEAGRHA 88

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||:|| |||| |||   | ||:|||||||||||||:|||||| ||||:|
Db         89 VEVWQGLALLSEAMLRSQALLANSSQLPETLQVHVDKAVSGLRSLTSLLRALGVQKEAVS 148

Qy        121 PPDAA-SAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              ||:|| |||||||: |||  ||||:||||||||||||||||| ||
Db        149 PPEAASSAAPLRTVAADTLCKLFRIYSNFLRGKLKLYTGEACRRGD 194


RESULT 9
I46578
erythropoietin - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C;Accession: I46578
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence homolo
A;Reference number: I46083; MUID:93372347; PMID:8364201
A;Accession: I46578
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-190 <WEN>
A;Cross-references: GB:L10607; NID:g164445; PIDN:AAA31029.1; PID:g164446
C;Superfamily: erythropoietin

  Query Match            80.1%;  Score 678;  DB 2;  Length 190;
  Best Local Similarity  82.0%;  Pred. No. 3.3e-58;
  Matches  137;  Conservative    7;  Mismatches   21;  Indels    2;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||:||||| || | |||| || :|||||||||||||||||||| |||
Db         23 APPRLICDSRVLERYILEAKEGENATMGCAESCSFSENITVPDTKVNFYAWKRMEVQQQA 82

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :|||||||||||:|:||||| ||||| | |||||||||||||||:||||||||||||
Db         83 MEVWQGLALLSEAILQGQALLANSSQPSEALQLHVDKAVSGLRSLTSLLRALGAQKEAIP 142

Qy        121 PPDA--ASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||  :|| ||||   ||  |||| |||||||| ||||||||  |
Db        143 LPDASPSSATPLRTFAVDTLCKLFRNYSNFLRGKLTLYTGEACRRRD 189


RESULT 10
I46199
erythropoietin - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 23-Jul-1999
C;Accession: I46199
R;Wen, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A;Title: Erythropoietin structure-function relationships: High degree of sequence hom
A;Reference number: I46083; MUID:93372347; PMID:8364201
A;Accession: I46199
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-175 <WEN>
A;Cross-references: GB:L13027; NID:g290087; PIDN:AAA30842.1; PID:g552347
C;Superfamily: erythropoietin

  Query Match            75.4%;  Score 638;  DB 2;  Length 175;
  Best Local Similarity  81.0%;  Pred. No. 2.2e-54;
  Matches  124;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||:|||:||||:| |||: || :|||||||||||| ||||:|||||
Db         23 APPRLICDSRVLERYILEAREAENVTMGCAQGCSFSENITVPDTKVNFYTWKRMDVGQQA 82

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :|||||||||||:||||||| |:||| |  |||||||| ||||||||||||||:|
Db         83 LEVWQGLALLSEAILRGQALLANASQPSETPQLHVDKAVSSLRSLTSLLRALGAQKEAMS 142

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 153
              |: || ||||| | ||  ||||:|||||||||
Db        143 LPEEASPAPLRTFTVDTLCKLFRIYSNFLRGKL 175


RESULT 11
G02729
thrombopoietin - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 05-Nov-1999
C;Accession: G02729
R;Im, S.
submitted to the EMBL Data Library, May 1996
A;Reference number: H01637
A;Accession: G02729
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-353 <IMX>
A;Cross-references: EMBL:U59493; NID:g1401245; PIDN:AAB03392.1; PID:g1401246
C;Genetics:
A;Gene: hTPO

  Query Match            10.6%;  Score 90;  DB 2;  Length 353;
  Best Local Similarity  26.3%;  Pred. No. 0.6;
  Matches   41;  Conservative   20;  Mismatches   75;  Indels   20;  Gaps    5;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
```

```
              |||   || ||| : | ::    :  : | |   |   : :|    :   ||      :|
Db         24 APP--ACDLRVLSKLLRDSHVLHSKLSQCPEVHPLPTPVLLPAVDFSLGEWKTQMEETKA 81

Qy         61 VEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEA 118
              ::   : || | |:  |||      |  .  |   |   :  |:||      || |
Db         82 QDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSEQVRLLLGALQSL------LGTQ--- 132

Qy        119 ISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKLK 154
                ||        ||        :|  : : ||||::
Db        133 -LPPQG------RTTAHKDPNAIFLSFQHLLRGKVR 161


RESULT 12
I80105
thrombopoietin precursor - human
N;Alternate names: c-MPL ligand; megakaryocyte growth and development factor precursor
C;Species: Homo sapiens (man)
C;Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 20-Jun-2000
C;Accession: I59281; I80105; S45331; S48740; I38672; I52610
R;Foster, D.C.; Sprecher, C.A.; Grant, F.J.; Kramer, J.M.; Kuijper, J.L.; Holly, R.D.; W
Proc. Natl. Acad. Sci. U.S.A. 91, 13023-13027, 1994
A;Title: Human thrombopoietin: gene structure, cDNA sequence, expression, and chromosoma
A;Reference number: I59281; MUID:95108091; PMID:7809166
A;Accession: I59281
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-353 <RE2>
A;Cross-references: GB:L36051; NID:g533214; PIDN:AAC37568.1; PID:g533215
A;Accession: I80105
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-353 <RES>
A;Cross-references: GB:L36052; NID:g533216; PIDN:AAC37566.1; PID:g533217
R;de Sauvage, F.J.; Hass, P.E.; Spencer, S.D.; Malloy, B.E.; Gurney, A.L.; Spencer, S.A.
D.V.; Eaton, D.L.
Nature 369, 533-538, 1994
A;Title: Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl ligand.
A;Reference number: S45331; MUID:94261202; PMID:8202154
A;Accession: S45331
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-353 <SAU>
A;Cross-references: GB:L33410; NID:g506826; PIDN:AAA59857.1; PID:g506827
R;Sohma, Y.; Akahori, H.; Seki, N.; Hori, T.; Ogami, K.; Kato, T.; Shimada, Y.; Kawamura
FEBS Lett. 353, 57-61, 1994
A;Title: Molecular cloning and chromosomal localization of the human thrombopoietin gene
A;Reference number: S48740; MUID:95010765; PMID:7926023
A;Accession: S48740
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-353 <SOH>
A;Cross-references: GB:D32046; NID:g577319; PIDN:BAA06807.1; PID:g577320
R;Bartley, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.S.;
Cell 77, 1117-1124, 1994
A;Title: Identification and cloning of a megakaryocyte growth and development factor tha
A;Reference number: A54463; MUID:94291201; PMID:8020099
A;Accession: I38672
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-112,'E',114-353 <RE3>
A;Cross-references: EMBL:U11025; NID:g511223; PIDN:AAA50553.1; PID:g558078
R;Gurney, A.L.; Kuang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.
Blood 85, 981-988, 1995
A;Title: Genomic structure, chromosomal localization, and conserved alternative splice f
A;Reference number: I52610; MUID:95152076; PMID:7849319
A;Accession: I52610
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-353 <RE4>
A;Cross-references: GB:S76771; NID:g914225; PIDN:AAB33390.1; PID:g914226
C;Genetics:
A;Gene: GDB:THPO; MGDF
A;Cross-references: GDB:374007; OMIM:600044
A;Map position: 3q26.3-3q27
A;Introns: 5/1; 47/3; 76/3; 132/3
C;Keywords: alternative splicing; cytokine; glycoprotein

  Query Match           10.5%;  Score 89;  DB 2;  Length 353;
  Best Local Similarity  26.3%;  Pred. No. 0.75;
  Matches   41;  Conservative   20;  Mismatches   75;  Indels   20;  Gaps    5;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||   || ||| : | ::    :  : | |   |   : :|    :   ||      :|
Db         24 APP--ACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPVLLPAVDFSLGEWKTQMEETKA 81

Qy         61 VEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEA 118
              ::   : || | |:  |||      |  .  |   |   :  |:||      || |
Db         82 QDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLLLGALQSL------LGTQ--- 132

Qy        119 ISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKLK 154
                ||        ||        :|  : : ||||::
Db        133 -LPPQG------RTTAHKDPNAIFLSFQHLLRGKVR 161


RESULT 13
AB0323
ribonucleoside-diphosphate reductase (EC 1.17.4.1) beta chain [imported] - Yersinia p
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C;Accession: AB0323
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AB0323
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-323 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC92889.1; PID:g15980631; GSPDB:GN00175
C;Genetics:
A;Gene: nrdF
C;Superfamily: ribonucleoside-diphosphate reductase beta
C;Keywords: oxidoreductase

  Query Match           10.4%;  Score 88;  DB 2;  Length 323;
  Best Local Similarity  25.2%;  Pred. No. 0.84;
  Matches   34;  Conservative   20;  Mismatches   59;  Indels   22;  Gaps    5;

Qy         38 NITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVD- 96
              |:  | |:::   | ::|   : :|||    | |   |  :  | |    | | |
Db          2 NVVKPITRISAINWNKIE-DDKDLEVWN--RLTSNFWLPEKVPLSNDIPSWATLTPHEQQ 58

Qy         97 ---KAVSGLRSLTTLLRALGAQ---KEAISPPDAASAAPLRTITADTFRKLFRVYSNFLR 150
                 :  :||  | |:   |||    |:||:| : |  : :  : |   |    ::|
Db         59 LTIRVFTGLTLLDTIQNTLGAPALIKDAITPHEEAIFSNISFMEAVHARSYSSIFSTL-- 116

Qy        151 GKLKLYTGEACRTGD 165
                        | | |
Db        117 ----------CLTSD 121


RESULT 14
AE0959
Solute binding receptor protein [imported] - Salmonella enterica subsp. enterica sero
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C;Accession: AE0959
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
```

```
A;Reference number: AB0502; PMID:11677608
A;Accession: AE0959
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-346 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD03169.1; PID:g16504804; GSPDB:GN00176
C;Genetics:
A;Gene: STY3952

  Query Match            10.3%;  Score 87.5;  DB 2;  Length 346;
  Best Local Similarity   26.7%;  Pred. No. 1;
  Matches   44;  Conservative   22;  Mismatches   48;  Indels   51;  Gaps    9;

Qy         10 RVLERYLLEAKEAENITTG--CAEHCSLNE--NITVPDTKVNFYAWKRMEVGQQAVEVWQ 65
              | | : :||     |:  |   |   :: |  |:| | | |:||            :|::
Db        217 RNLLQEMLERHPDANVVAGSAIAAEAAMGEGRNLTTPLTIVSFYL---------THQVYR 267

Qy         66 GLALLSEAVLRGQALLVNSSQ-PWEPLQLHVDKAVSGLRSLTTLLRALGAQ--KEAISPP 122
              ||        ||  |:  | |  |:          |  ::|  :: |  |   | ||||
Db        268 GLK-------RGHILMALSDQMAWQ----------GELAITQSIKVLQGQPVPENISPP 309

Qy        123 -------DAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEA 160
                   :| ||   |:::   || ::          |  || ||
Db        310 VLILTHNNADSARVRRSLSPPGFRPVY----------LYQYTSEA 344

RESULT 15
A55530
megakaryocyte growth and development factor, long form - human
N;Alternate names: MPL ligand, long form
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-May-1999
C;Accession: A55530
R;Chang, M.; McNinch, J.; Basu, R.; Shutter, J.; Hsu, R.; Perkins, C.; Mar, V.; Suggs, S
J. Biol. Chem. 270, 511-514, 1995
A;Title: Cloning and characterization of the human megakoryocyte growth and development
A;Reference number: A55530; MUID:95122483; PMID:7822271
A;Accession: A55530
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-286 <CHA>
A;Cross-references: GB:U17071
C;Genetics:
A;Gene: MGDF
A;Map position: 3q26.3
C;Keywords: alternative splicing; cytokine

  Query Match            10.2%;  Score 86;  DB 2;  Length 286;
  Best Local Similarity   26.6%;  Pred. No. 1.1;
  Matches   41;  Conservative   18;  Mismatches   75;  Indels   20;  Gaps    5;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||   || ||| : | ::    :  : | |   |   : :|    :   ||      :|
Db         24 APP--ACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPVLLPAVDFSLGEWKTQMEETKA 81

Qy         61 VEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEA 118
               ::   : || | |:  |||     |    |   |   :  |:||      || |
Db         82 QDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLLLGALQSL------LGTQ--- 132

Qy        119 ISPPDAASAAPLRTITADTFRKLFRVYSNFLRGK 152
                ||        ||        :|  : : ||||
Db        133 -LPPQG------RTTAHKDPNAIFLSFQHLLRGK 159


Search completed: June 13, 2003, 15:04:42
Job time : 33 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:03:00 ; Search time 12 Seconds
(without alignments)
570.300 Million cell updates/sec

Title: US-09-830-964-1
Perfect score: 846
Sequence: 1 APPRLICDSRVLERYLLEAK.....SNFLRGKLKLYTGEACRTGD 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	846	100.0	193	1	EPO_HUMAN	P01588 homo sapien
2	764.5	90.4	192	1	EPO_MACFA	P07865 macaca fasc
3	759.5	89.8	192	1	EPO_MACMU	Q28513 macaca mula
4	706	83.5	192	1	EPO_FELCA	P33708 felis silve
5	701	82.9	192	1	EPO_RAT	P29676 rattus norv
6	692.5	81.9	192	1	EPO_BOVIN	P48617 bos taurus
7	689	81.4	192	1	EPO_MOUSE	P07321 mus musculu
8	685.5	81.0	194	1	EPO_SHEEP	P33709 ovis aries
9	678	80.1	190	1	EPO_PIG	P49157 sus scrofa
10	638	75.4	175	1	EPO_CANFA	P33707 canis famil
11	109	12.9	352	1	TPO_CANFA	P42705 canis famil
12	89	10.5	353	1	TPO_HUMAN	P40225 homo sapien
13	80.5	9.5	3033	1	POLG_HCVJ8	P26661 h genome po
14	78.5	9.3	897	1	EP15_MOUSE	P42567 mus musculu
15	78	9.2	747	1	ATCS_SYNP7	P37279 synechococc
16	77	9.1	548	1	CH60_BUCMP	O51832 buchnera ap
17	76.5	9.0	386	1	CTBP_DROME	O46036 drosophila
18	76	9.0	551	1	CH60_BUCAP	Q59177 buchnera ap
19	74.5	8.8	552	1	CH60_PSEST	O33500 pseudomonas
20	74.5	8.8	809	1	HRPB_ECOLI	P37024 escherichia
21	74	8.7	326	1	TPO_RAT	P49745 rattus norv
22	74	8.7	1564	1	PDRA_YEAST	P51533 saccharomyc
23	73.5	8.7	830	1	GCL2_MOUSE	Q9jhd2 mus musculu
24	73.5	8.7	837	1	GCL2_HUMAN	Q92830 homo sapien
25	73	8.6	263	1	YH25_DEIRA	O83030 deinococcus
26	73	8.6	1089	1	IMB3_YEAST	P32337 saccharomyc
27	72.5	8.6	762	1	SLAP_ACEKI	P22258 acetogenium
28	72	8.5	356	1	TPO_MOUSE	P40226 mus musculu
29	72	8.5	548	1	CH60_BUCAI	P25750 buchnera ap
30	70.5	8.3	217	1	YNJD_ECOLI	P76909 escherichia
31	70.5	8.3	381	1	MODD_MYCAV	Q48919 mycobacteri
32	70.5	8.3	543	1	CH60_BARBA	P35635 bartonella
33	70.5	8.3	896	1	EP15_HUMAN	P42566 homo sapien
34	70	8.3	319	1	RIR4_SALTY	P17424 salmonella
35	70	8.3	547	1	CH60_LEGPN	P26878 legionella
36	69.5	8.2	907	1	GACS_PSESY	P48027 pseudomonas
37	69.5	8.2	1327	1	TNK1_HUMAN	O95271 homo sapien
38	69	8.2	353	1	NADA_YERPE	Q8zgy8 yersinia pe
39	69	8.2	544	1	CH60_FRATU	P94798 francisella
40	69	8.2	552	1	CH60_COXBU	P19421 coxiella bu
41	69	8.2	1128	1	DNBI_HSVSA	P24910 herpesvirus
42	68.5	8.1	360	1	CYS2_MAIZE	Q10717 zea mays (m
43	68.5	8.1	418	1	GUN2_TRIRE	P07982 trichoderma
44	68.5	8.1	778	1	RGL2_MOUSE	Q61193 mus musculu
45	68	8.0	366	1	QUEA_SYNY3	Q55850 synechocyst

ALIGNMENTS

RESULT 1
EPO_HUMAN
ID EPO_HUMAN STANDARD; PRT; 193 AA.
AC P01588; Q9UHA0; Q9UEZ5; Q9UDZ0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor (Epoetin).
GN EPO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85137899; PubMed=3838366;
RA Jacobs K., Shoemaker C., Rudersdorf R., Neill S.D., Kaufman R.J.,
RA Mufson A., Seehra J., Jones S.S., Hewick R., Fritsch E.F.,
RA Kawakita M., Shimizu T., Miyake T.;
RT "Isolation and characterization of genomic and cDNA clones of human
RT erythropoietin.";
RL Nature 313:806-810(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067948; PubMed=3865178;
RA Lin F.-K., Suggs S., Lin C.-H., Browne J.K., Smalling R., Egrie J.C.,
RA Chen K.K., Fox G.M., Martin F., Stabinsky Z., Badrawi S.M., Lai P.-H.,
RA Goldwasser E.;
RT "Cloning and expression of the human erythropoietin gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7580-7584(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99018118; PubMed=9799793;
RA Gloeckner G., Scherer S., Schattevoy R., Boright A., Weber J.,
RA Tsui L.-C., Rosenthal A.;
RT "Large-scale sequencing of two regions in human chromosome 7q22:
RT analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci
RT reveals 17 genes.";
RL Genome Res. 8:1060-1073(1998).
RN [4]
RP SEQUENCE FROM N.A.
RA Rupert J.L., Hochachka P.W.;
RT "Erythropoietin gene sequence in the Quechua, a high altitude native
RT population.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE OF 58-193 FROM N.A., AND VARIANTS HEPATOCELLULAR CARCINOMA.
RX MEDLINE=93384593; PubMed=8396923;
RA Funakoshi A., Muta H., Baba T., Shimizu S.;
RT "Gene expression of mutant erythropoietin in hepatocellular
RT carcinoma.";
RL Biochem. Biophys. Res. Commun. 195:717-722(1993).
RN [6]
RP SEQUENCE OF 28-193, AND DISULFIDE BONDS.
RC TISSUE=Urine;
RX MEDLINE=86140080; PubMed=3949763;

```
RA   Lai P.H., Everett R., Wang F.F., Arakawa T., Goldwasser E.;
RT   "Structural characterization of human erythropoietin.";
RL   J. Biol. Chem. 261:3116-3121(1986).
RN   [7]
RP   PRELIMINARY SEQUENCE OF 28-57.
RX   MEDLINE=84135751; PubMed=6698989;
RA   Yanagawa S., Hirade K., Ohnota H., Sasaki R., Chiba H., Ueda M.,
RA   Goto M.;
RT   "Isolation of human erythropoietin with monoclonal antibodies.";
RL   J. Biol. Chem. 259:2707-2710(1984).
RN   [8]
RP   STRUCTURE OF CARBOHYDRATES.
RX   MEDLINE=88153657; PubMed=3346214;
RA   Takeuchi M., Takasaki S., Miyazaki H., Kato T., Hoshi S., Kochibe N.,
RA   Kobata A.;
RT   "Comparative study of the asparagine-linked sugar chains of human
RT   erythropoietins purified from urine and the culture medium of
RT   recombinant Chinese hamster ovary cells.";
RL   J. Biol. Chem. 263:3657-3663(1988).
RN   [9]
RP   STRUCTURE OF CARBOHYDRATES.
RX   MEDLINE=89118279; PubMed=3219367;
RA   Sasaki H., Ochi N., Dell A., Fukuda M.;
RT   "Site-specific glycosylation of human recombinant erythropoietin:
RT   analysis of glycopeptides or peptides at each glycosylation site by
RT   fast atom bombardment mass spectrometry.";
RL   Biochemistry 27:8618-8626(1988).
RN   [10]
RP   STRUCTURE OF CARBOHYDRATES.
RX   MEDLINE=92314463; PubMed=1820196;
RA   Takeuchi M., Kobata A.;
RT   "Structures and functional roles of the sugar chains of human
RT   erythropoietins.";
RL   Glycobiology 1:337-346(1991).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=98445092; PubMed=9774108;
RA   Syed R.S., Reid S.W., Li C., Cheetham J.C., Aoki K.H., Liu B.,
RA   Zhan H., Osslund T.D., Chirino A.J., Zhang J., Finer-Moore J.,
RA   Elliott S., Sitney K., Katz B.A., Matthews D.J., Wendoloski J.J.,
RA   Egrie J., Stroud R.M.;
RT   "Efficiency of signalling through cytokine receptors depends
RT   critically on receptor orientation.";
RL   Nature 395:511-516(1998).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- PHARMACEUTICAL: Used for the treatment of anemia. Available under
CC       the names Epogen (Amgen), Epogin (Chugai), Epomax (Elanex), Eprex
CC       (Janssen-Cilag), NeoRecormon or Recormon (Roche), and Procrit
CC       (Ortho Biotech). Variations in the glycosylation pattern of EPO
CC       distinguishes these products. Epogen, Epogin, Eprex and Procrit
CC       are generically known as epoetin alfa, NeoRecormon and Recormon as
CC       epoetin beta and Epomax as epoetin omega.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   -!- DATABASE: NAME=R&D Systems' cytokine source book: EPO;
CC       WWW="http://www.rndsystems.com/asp/g_sitebuilder.asp?bodyId=197".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X02158; CAA26095.1; -.
DR   EMBL; X02157; CAA26094.1; -.
DR   EMBL; M11319; AAA52400.1; -.
DR   EMBL; AF053356; AAC78791.1; -.
DR   EMBL; AF202308; AAF23132.1; -.
DR   EMBL; AF202306; AAF23132.1; JOINED.
DR   EMBL; AF202307; AAF23132.1; JOINED.
DR   EMBL; AF202310; AAF23133.1; -.
DR   EMBL; AF202309; AAF23133.1; JOINED.
DR   EMBL; AF202311; AAF17572.1; -.
DR   EMBL; AF202314; AAF23134.1; -.
DR   EMBL; AF202312; AAF23134.1; JOINED.
DR   EMBL; AF202313; AAF23134.1; JOINED.
DR   EMBL; S65458; AAD13964.1; -.
DR   PIR; A01855; ZUHU.
DR   PIR; A25384; A25384.
DR   PIR; A24744; A24744.
DR   PIR; A22210; A22210.
DR   PDB; 1EER; 01-OCT-99.
DR   PDB; 1CN4; 11-AUG-99.
DR   GlycoSuiteDB; P01588; -.
DR   Genew; HGNC:3415; EPO.
DR   MIM; 133170; -.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal; Pharmaceutical;
KW   3D-structure.
FT   SIGNAL       1    27
FT   CHAIN       28   193       ERYTHROPOIETIN.
FT   PROPEP     190   193       MAY BE REMOVED IN PROCESSED PROTEIN.
FT   DISULFID    34   188
FT   DISULFID    56    60
FT   CARBOHYD    51    51       N-LINKED (GLCNAC. . .).
FT                               /FTId=CAR_000052.
FT   CARBOHYD    65    65       N-LINKED (GLCNAC. . .).
FT                               /FTId=CAR_000166.
FT   CARBOHYD   110   110       N-LINKED (GLCNAC. . .).
FT                               /FTId=CAR_000192.
FT   CARBOHYD   153   153       O-LINKED (GALNAC. . .).
FT   VARIANT    131   132       SL -> NF (IN AN HEPATOCELLULAR
FT                               CARCINOMA).
FT                               /FTId=VAR_009870.
FT   VARIANT    149   149       P -> Q (IN AN HEPATOCELLULAR CARCINOMA).
FT                               /FTId=VAR_009871.
FT   CONFLICT    40    40       E -> Q (IN REF. 1; CAA26095).
FT   CONFLICT    85    85       Q -> QQ (IN REF. 5).
FT   CONFLICT   140   140       G -> R (IN REF. 1; CAA26095).
SQ   SEQUENCE   193 AA;  21306 MW;  C91F0E4C26A52033 CRC64;

  Query Match             100.0%;  Score 846;  DB 1;  Length 193;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-76;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 147

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db        148 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 192


RESULT 2
EPO_MACFA
ID   EPO_MACFA      STANDARD;      PRT;   192 AA.
AC   P07865;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
```

```
GN   EPO.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87055236; PubMed=2877922;
RA   Lin F.-K., Lin C.-H., Lai P.-H., Browne J.K., Egrie J.C., Smalling R.,
RA   Fox G.M., Chen K.K., Castro M., Suggs S.;
RT   "Monkey erythropoietin gene: cloning, expression and comparison with
RT   the human erythropoietin gene.";
RL   Gene 44:201-209(1986).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M18189; AAA36841.1; -.
DR   PIR; JQ0173; JQ0173.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   SIGNAL       1     27       BY SIMILARITY.
FT   CHAIN       28    192       ERYTHROPOIETIN.
FT   DISULFID    34    187       BY SIMILARITY.
FT   DISULFID    56     60       BY SIMILARITY.
FT   CARBOHYD    51     51       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD    65     65       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD   110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD   152    152       O-LINKED (GALNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   192 AA;  21113 MW;  E8A900F442AD4522 CRC64;

  Query Match             90.4%;  Score 764.5;  DB 1;  Length 192;
  Best Local Similarity   91.5%;  Pred. No. 2.2e-68;
  Matches  151;  Conservative    7;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||||||||||||||||||||:| ||:| ||||||||||||||||||||||||||||
Db         28 APPRLICDSRVLERYLLEAKEAENVTMGCSESCSLNENITVPDTKVNFYAWKRMEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              |||||||||||||||||||:| ||||:||||||:|||:|||||:|||||||||| ||||
Db         88 VEVWQGLALLSEAVLRGQAVLANSSQPFEPLQLHMDKAISGLRSITTLLRALGAQ-EAIS 146

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
               |||||||||||||||| |||||||||||||||||||||| ||
Db        147 LPDAASAAPLRTITADTFCKLFRVYSNFLRGKLKLYTGEACRRGD 191


RESULT 3
EPO_MACMU
ID   EPO_MACMU      STANDARD;      PRT;   192 AA.
AC   Q28513;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=93372347; PubMed=8364201;
RA   Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA   Czelusniak J., Goodman M., Bunn H.F.;
RT   "Erythropoietin structure-function relationships: high degree of
RT   sequence homology among mammals.";
RL   Blood 82:1507-1516(1993).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L10609; AAA36842.1; -.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   SIGNAL       1     27       BY SIMILARITY.
FT   CHAIN       28    192       ERYTHROPOIETIN.
FT   DISULFID    34    187       BY SIMILARITY.
FT   DISULFID    56     60       BY SIMILARITY.
FT   CARBOHYD    51     51       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD    65     65       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD   110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD   152    152       O-LINKED (GALNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   192 AA;  21081 MW;  275560A264628CD1 CRC64;

  Query Match             89.8%;  Score 759.5;  DB 1;  Length 192;
  Best Local Similarity   90.3%;  Pred. No. 6.7e-68;
  Matches  149;  Conservative    9;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||:|||||||||||||||:| ||:| |||||||||||||||||||:||||||
Db         28 APPRLVCDSRVLERYLLEAKEAENVTMGCSESCSLNENITVPDTKVNFYAWKRIEVGQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              |||||||||||||||||:| |||||:||||||:|||:|||||:|||||||||| ||||
Db         88 VEVWQGLALLSEAVLRGQAVLANSSQPFEPLQLHMDKAISGLRSITTLLRALGAQ-EAIS 146

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
               |||||||||||||||| |||||||||||||||||||||| ||
Db        147 LPDAASAAPLRTITADTFCKLFRVYSNFLRGKLKLYTGEACRRGD 191


RESULT 4
EPO_FELCA
ID   EPO_FELCA      STANDARD;      PRT;   192 AA.
AC   P33708;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
```

```
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX   NCBI_TaxID=9685;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Goodman R.E., Bell R.G.;
RL   Submitted (NOV-1993) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 5-192 FROM N.A.
RX   MEDLINE=93372347; PubMed=8364201;
RA   Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA   Czelusniak J., Goodman M., Bunn H.F.;
RT   "Erythropoietin structure-function relationships: high degree of
RT   sequence homology among mammals.";
RL   Blood 82:1507-1516(1993).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U00685; AAA18282.1; -.
DR   EMBL; L10606; AAA30807.1; -.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    192       ERYTHROPOIETIN.
FT   DISULFID    33    187       BY SIMILARITY.
FT   DISULFID    55     59       BY SIMILARITY.
FT   CARBOHYD    50     50       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    64     64       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   109    109       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    44     44       G -> E (IN REF. 2).
SQ   SEQUENCE   192 AA;  20914 MW;  61C5EA0F5E937293 CRC64;

  Query Match           83.5%;  Score 706;  DB 1;  Length 192;
  Best Local Similarity  83.6%;  Pred. No. 1.3e-62;
  Matches  138;  Conservative    9;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||:| |:||||:| |||| || :|||||||||||| ||||:|||||
Db         27 APPRLICDSRVLERYILGAREAENVTMGCAEGCSFSENITVPDTKVNFYTWKRMDVGQQA 86

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              |||||||||||||:||||||| ||||| | |||||||||| |||||:|||||||||||| |
Db         87 VEVWQGLALLSEAILRGQALLANSSQPSETLQLHVDKAVSSLRSLTSLLRALGAQKEATS 146

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
               |:| ||||||| | ||  |||||:|||||||| |||||||| ||
Db        147 LPEATSAAPLRTFTVDTLCKLFRIYSNFLRGKLTLYTGEACRRGD 191


RESULT 5
```

```
EPO_RAT
ID   EPO_RAT        STANDARD;      PRT;   192 AA.
AC   P29676; P70504;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Kidney;
RX   MEDLINE=93042015; PubMed=1420369;
RA   Nagao M., Suga H., Okano M., Masuda S., Narita H., Ikura K.,
RA   Sasaki R.;
RT   "Nucleotide sequence of rat erythropoietin.";
RL   Biochim. Biophys. Acta 1171:99-102(1992).
RN   [2]
RP   SEQUENCE OF 4-192 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX   MEDLINE=93372347; PubMed=8364201;
RA   Wen D., Boissel J.P.R., Tracy T.E., Mulcahy L.S., Czelusniak J.,
RA   Goodman M., Bunn H.F.;
RT   "Erythropoietin structure-function relationships: high degree of
RT   sequence homology among mammals.";
RL   Blood 82:1507-1516(1993).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10763; BAA01593.1; -.
DR   EMBL; L10608; AAA41126.1; -.
DR   PIR; S28148; S28148.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    192       ERYTHROPOIETIN.
FT   DISULFID    33    187       BY SIMILARITY.
FT   CARBOHYD    50     50       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD    64     64       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD   109    109       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   192 AA;  21286 MW;  3EA632737E7D2443 CRC64;

  Query Match           82.9%;  Score 701;  DB 1;  Length 192;
  Best Local Similarity  82.4%;  Pred. No. 4e-62;
  Matches  136;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||:|||||||:| ||||   |:||||||||||||||||:| :||
Db         27 APPRLICDSRVLERYILEAKEAENVTMGCAEGPRLSENITVPDTKVNFYAWKRMKVEEQA 86

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              |||||||:|||||:|: |||  ||||| | ||||:|||:|||||||:||| |||||| :|
```

```
Db         87 VEVWQGLSLLSEAILQAQALQANSSQPPESLQLHIDKAISGLRSLTSLLRVLGAQKELMS 146

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              ||||  ||||||:||||| |||||||||||||||||||||||| ||
Db        147 PPDATQAAPLRTLTADTFCKLFRVYSNFLRGKLKLYTGEACRRGD 191


RESULT 6
EPO_BOVIN
ID   EPO_BOVIN      STANDARD;      PRT;   192 AA.
AC   P48617;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Boran; TISSUE=Kidney;
RX   MEDLINE=96257233; PubMed=8666286;
RA   Suliman H.B., Majiwa P.A.O., Feldman B.F., Mertens B.,
RA   Logan-Henfrey L.L.;
RT   "Cloning of a cDNA encoding bovine erythropoietin and analysis of its
RT   transcription in selected tissues.";
RL   Gene 171:275-280(1996).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L41354; AAB41268.1; -.
DR   EMBL; U44762; AAA86653.1; -.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    192       ERYTHROPOIETIN.
FT   DISULFID     32    187       BY SIMILARITY.
FT   DISULFID     54     58       BY SIMILARITY.
FT   CARBOHYD     49     49       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    108    108       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   192 AA;  21075 MW;  DBC419022F7B483A CRC64;

  Query Match             81.9%;  Score 692.5;  DB 1;  Length 192;
  Best Local Similarity   83.1%;  Pred. No. 2.8e-61;
  Matches  138;  Conservative    8;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              || |||||||||||:|||:|||| | |||| || ||||||||||||||||||| |||
Db         26 APARLICDSRVLERYILEAREAENATMGCAEGCSFNENITVPDTKVNFYAWKRMEVQQQA 85

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :||||||||||||:||||||| |:||| | |:||||||||||||:|||||||||||
Db         86 LEVWQGLALLSEAILRGQALLANASQPCEALRLHVDKAVSGLRSLTSLLRALGAQKEAIS 145

Qy        121 PPDAA-SAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
               |||  |||||| | |   ||||:||||||||| |||||||| ||
Db        146 LPDATPSAAPLRAFTVDALSKLFRIYSNFLRGKLTLYTGEACRRGD 191


RESULT 7
EPO_MOUSE
ID   EPO_MOUSE      STANDARD;      PRT;   192 AA.
AC   P07321;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87039105; PubMed=3773894;
RA   Shoemaker C.B., Mitsock L.D.;
RT   "Murine erythropoietin gene: cloning, expression, and human gene
RT   homology.";
RL   Mol. Cell. Biol. 6:849-858(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87039104; PubMed=3022133;
RA   McDonald J.D., Lin F.-K., Goldwasser E.;
RT   "Cloning, sequencing, and evolutionary analysis of the mouse
RT   erythropoietin gene.";
RL   Mol. Cell. Biol. 6:842-848(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv;
RX   MEDLINE=21138439; PubMed=11239002;
RA   Wilson M.D., Riemer C., Martindale D.W., Schnupf P., Boright A.P.,
RA   Cheung T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,
RA   Miller W., Koop B.F.;
RT   "Comparative analysis of the gene-dense ACHE/TFR2 region on human
RT   chromosome 7q22 with the orthologous region on mouse chromosome 5.";
RL   Nucleic Acids Res. 29:1352-1365(2001).
RN   [4]
RP   SEQUENCE OF 1-52 FROM N.A.
RC   STRAIN=ICFW;
RX   MEDLINE=98030528; PubMed=9365246;
RA   Chretien S., Duprez V., Maouche L., Gisselbrecht S., Mayeux P.,
RA   Lacombe C.;
RT   "Abnormal erythropoietin (Epo) gene expression in the murine
RT   erythroleukemia IW32 cells results from a rearrangement between the
RT   G- protein beta2 subunit gene and the Epo gene.";
RL   Oncogene 15:1995-1999(1997).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M12482; AAA37568.1; -.
DR   EMBL; M12930; AAA37570.1; -.
```

```
DR   EMBL; AF312033; AAK28825.1; -.
DR   EMBL; Y11971; CAA72707.1; -.
DR   PIR; A24901; A24901.
DR   PIR; A24902; A24902.
DR   HSSP; P01588; 1CN4.
DR   MGD; MGI:95407; Epo.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   SIGNAL       1     26
FT   CHAIN       27    192       ERYTHROPOIETIN.
FT   DISULFID    33    187       BY SIMILARITY.
FT   CARBOHYD    50     50       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD    64     64       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD   109 .  109       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   192 AA;  21365 MW;  65F94E214E0DEF2E CRC64;

  Query Match             81.4%;  Score 689;  DB 1;  Length 192;
  Best Local Similarity   80.0%;  Pred. No. 6.2e-61;
  Matches  132;  Conservative   14;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||||||||||||:||||||||:| ||||   |:||||||||||||||||||| :||
Db         27 APPRLICDSRVLERYILEAKEAENVTMGCAEGPRLSENITVPDTKVNFYAWKRMEVEEQA 86

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :||||||:|||||:|: |||| |.||| | ||||:|||:|||||||:||| |||||| :|
Db         87 IEVWQGLSLLSEAILQAQALLANSSQPPETLQLHIDKAISGLRSLTSLLRVLGAQKELMS 146

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              |||    |||||:| ||| |||||:|||||||||||| || ||
Db        147 PPDTTPPAPLRTLTVDTFCKLFRVYANFLRGKLKLYTGEVCRRGD 191


RESULT 8
EPO_SHEEP
ID   EPO_SHEEP      STANDARD;      PRT;   194 AA.
AC   P33709; Q28572;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=93351736; PubMed=8349021;
RA   Fu P., Evans B., Lim G.B., Moritz K., Wintour M.E.;
RT   "The sheep erythropoietin gene: molecular cloning and effect of
RT   hemorrhage on plasma erythropoietin and renal/liver messenger RNA in
RT   adult sheep.";
RL   Mol. Cell. Endocrinol. 93:107-116(1993).
RN   [2]
RP   SEQUENCE OF 4-194 FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=93372347; PubMed=8364201;
RA   Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA   Czelusniak J., Goodman M., Bunn H.F.;
RT   "Erythropoietin structure-function relationships: high degree of
RT   sequence homology among mammals.";
RL   Blood 82:1507-1516(1993).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z24681; CAA80848.1; -.
DR   EMBL; L10610; AAA31518.1; -.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   SIGNAL       1     27       BY SIMILARITY.
FT   CHAIN       28    194       ERYTHROPOIETIN.
FT   DISULFID    34    189       BY SIMILARITY.
FT   DISULFID    56     60       BY SIMILARITY.
FT   CARBOHYD    51     51       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    65     65       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    16     16       F -> L (IN REF. 2).
FT   CONFLICT   108    108       L -> P (IN REF. 2).
SQ   SEQUENCE   194 AA;  21335 MW;  C025AAB05281131A9 CRC64;

  Query Match             81.0%;  Score 685.5;  DB 1;  Length 194;
  Best Local Similarity   81.9%;  Pred. No. 1.4e-60;
  Matches  136;  Conservative    9;  Mismatches   20;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||||||||||||||:|||:|||| | |||| || :|||||||||||||||||||| |||
Db         28 APPRLICDSRVLERYILEAREAENATMGCAEGCSFSENITVPDTKVNFYAWKRMEVQQQA 87

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :||||||||||||: |||||| |:||| | |:|||||||||||||:||||||||||||
Db         88 LEVWQGLALLSEAIFRGQALLANASQPCEALRLHVDKAVSGLRSLTSLLRALGAQKEAIP 147

Qy        121 PPDAA-SAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
                 |||  |||||  | |   ||||:||||||||| |||||||| ||
Db        148 LPDATPSAAPLRIFTVDALSKLFRIYSNFLRGKLTLYTGEACRRGD 193


RESULT 9
EPO_PIG
ID   EPO_PIG        STANDARD;      PRT;   190 AA.
AC   P49157;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor (Fragment).
GN   EPO.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=93372347; PubMed=8364201;
RA   Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA   Czelusniak J., Goodman M., Bunn H.F.;
RT   "Erythropoietin structure-function relationships: high degree of
RT   sequence homology among mammals.";
RL   Blood 82:1507-1516(1993).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
```

```
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L10607; AAA31029.1; -.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     22       POTENTIAL.
ET   CHAIN        23    190       ERYTHROPOIETIN.
FT   DISULFID     29    185       BY SIMILARITY.
FT   DISULFID     51     55       BY SIMILARITY.
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    105    105       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   190 AA;  20888 MW;  A75BD6CCE5077E2A CRC64;

  Query Match             80.1%;  Score 678;  DB 1;  Length 190;
  Best Local Similarity   82.0%;  Pred. No. 7.5e-60;
  Matches  137;  Conservative    7;  Mismatches   21;  Indels    2;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||:|||| || | |||| || :|||||||||||||||||||| |||
Db         23 APPRLICDSRVLERYILEAKEGENATMGCAESCSFSENITVPDTKVNFYAWKRMEVQQQA 82

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :|||||||||||:|:|||| |||| | |||||||||||||||:||||||||||||
Db         83 MEVWQGLALLSEAILQGQALLANSSQPSEALQLHVDKAVSGLRSLTSLLRALGAQKEAIP 142

Qy        121 PPDA--ASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
                 |||  :|| ||||   ||  |||| ||||||||| ||||||||  |
Db        143 LPDASPSSATPLRTFAVDTLCKLFRNYSNFLRGKLTLYTGEACRRRD 189


RESULT 10
EPO_CANFA
ID   EPO_CANFA      STANDARD;      PRT;   175 AA.
AC   P33707;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Erythropoietin precursor (Fragment).
GN   EPO.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93372347; PubMed=8364201;
RA   Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA   Czelusniak J., Goodman M., Bunn H.F.;
RT   "Erythropoietin structure-function relationships: high degree of
RT   sequence homology among mammals.";
RL   Blood 82:1507-1516(1993).
CC   -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC       REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC       PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC       AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L13027; AAA30842.1; -.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     22       BY SIMILARITY.
FT   CHAIN        23   >175       ERYTHROPOIETIN.
FT   DISULFID     29   >175       BY SIMILARITY.
FT   DISULFID     51     55       BY SIMILARITY.
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    105    105       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     175    175
SQ   SEQUENCE   175 AA;  19193 MW;  B504F8DE86676BF4 CRC64;

  Query Match             75.4%;  Score 638;  DB 1;  Length 175;
  Best Local Similarity   81.0%;  Pred. No. 6e-56;
  Matches  124;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||:|||:||||:| |||: || :|||||||||||| ||||:|||||
Db         23 APPRLICDSRVLERYILEAREAENVTMGCAQGCSFSENITVPDTKVNFYTWKRMDVGQQA 82

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :|||||||||||:|||||| |:||| |  ||||||||| |||||:||||||||||:|
Db         83 LEVWQGLALLSEAILRGQALLANASQPSETPQLHVDKAVSSLRSLTSLLRALGAQKEAMS 142

Qy        121 PPDAASAAPLRTITADTFRKLFRVYSNFLRGKL 153
              |: || ||||| | ||  ||||:|||||||||
Db        143 LPEEASPAPLRTFTVDTLCKLFRIYSNFLRGKL 175


RESULT 11
TPO_CANFA
ID   TPO_CANFA      STANDARD;      PRT;   352 AA.
AC   P42705;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
DE   (C-MPL ligand) (ML) (Megakaryocyte growth and development factor)
DE   (MGDF).
GN   THPO OR TPO.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 24-44.
RC   TISSUE=Kidney;
RX   MEDLINE=94291201; PubMed=8020099;
RA   Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
RA   Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
RA   Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,
RA   Clogson C., Hsu E., Hokom M.M., Hornkohl A., Choi E., Pangelinan M.,
RA   Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
RA   Shutter J., Chute H., Basu R., Selander L., Trollinger D., Sieu L.,
RA   Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
RA   Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,
```

```
RA   Pacifici R., Ponting I., Saris C., Wen D., Yung Y.P., Lin H.,
RA   Bosselman R.A.;
RT   "Identification and cloning of a megakaryocyte growth and development
RT   factor that is a ligand for the cytokine receptor Mpl.";
RL   Cell 77:1117-1117(1994).
CC   -!- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
CC       AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
CC       CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
CC       MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
CC       TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003978; thrombopoeitin.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR01485; THROMBOPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Cytokine; Glycoprotein; Hormone; Signal.
FT   SIGNAL       1     23
FT   CHAIN       24    352       THROMBOPOIETIN.
FT   DISULFID    28    172       POTENTIAL.
FT   DISULFID    50    106       POTENTIAL.
FT   CARBOHYD   185    185       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   206    206       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   255    255       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   332    332       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   347    347       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   352 AA;  37641 MW;  024F3B41B061FBD8 CRC64;

  Query Match           12.9%;  Score 109;  DB 1;  Length 352;
  Best Local Similarity  24.0%;  Pred. No. 0.0025;
  Matches   40;  Conservative   23;  Mismatches   62;  Indels   42;  Gaps    5;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||   || |:| : | ::    :  : | :   |:  : :|    :   ||  :  :|
Db         24 APP--ACDPRLLNKMLRDSHVLHSRLSQCPDIYPLSTPVLLPAVDFSLGEWKTQKEQTKA 81

Qy         61 VEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEA 118
               :||  :||| : ||  |||                    |   |::||  |  |
Db         82 QDVWGAVALLLDGVLAARGQL--------------------GPSCLSSLLGQLSGQVRL 120

Qy        119 I----------SPPDAASAAPLRTITADTFRKLFRVYSNFLRGKLK 154
              :          ||       || |      :|  :   ||||::
Db        121 LLGALQGLLGTQLPPQG------RTTTHKDPNAIFLSFQQLLRGKVR 161


RESULT 12
TPO_HUMAN
ID   TPO_HUMAN      STANDARD;      PRT;   353 AA.
AC   P40225; Q13020; Q15790; Q15791; Q15792;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
DE   (Myeloproliferative leukemia virus oncogene ligand) (C-mpl ligand)
DE   (ML) (Megakaryocyte growth and development factor) (MGDF).
GN   THPO.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Fetal liver;
RX   MEDLINE=94261202; PubMed=8202154;
RA   de Sauvage F.J., Hass P.E., Spencer S.D., Malloy B.E., Gurney A.L.,
RA   Spencer S.A., Darbonne W.C., Henzel W.J., Wong S.C., Kuang W.-J.,
RA   Oles K.J., Hultgren B., Solberg L.A. Jr., Goeddel D.V., Eaton D.L.;
RT   "Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl
RT   ligand.";
RL   Nature 369:533-538(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Fetal liver;
RX   MEDLINE=94291201; PubMed=8020099;
RA   Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
RA   Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
RA   Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,
RA   Clogson C., Hsu E., Hokom M.M., Hornkohl A., Choi E., Pangelinan M.,
RA   Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
RA   Shutter J., Chute H., Basu R., Selander L., Trollinger D., Sieu L.,
RA   Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
RA   Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,
RA   Pacifici R., Ponting I., Saris C., Wen D., Yung Y.P., Lin H.,
RA   Bosselman R.A.;
RT   "Identification and cloning of a megakaryocyte growth and development
RT   factor that is a ligand for the cytokine receptor Mpl.";
RL   Cell 77:1117-1124(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=95108091; PubMed=7809166;
RA   Foster D.C., Sprecher C.A., Grant F.J., Kramer J.M., Kuijper J.L.,
RA   Holly R.D., Whitmore T.E., Heipel M.D., Bell L.A.N., Ching A.F.,
RA   McGrane V., Hart C., O'Hara P.J., Lok S.;
RT   "Human thrombopoietin: gene structure, cDNA sequence, expression, and
RT   chromosomal localization.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:13023-13027(1994).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=95010765; PubMed=7926023;
RA   Sohma Y., Akahori H., Seki N., Hori T.-A., Ogami K., Kawamura K.,
RA   Miyazaki H.;
RT   "Molecular cloning and chromosomal localization of the human
RT   thrombopoietin gene.";
RL   FEBS Lett. 353:57-61(1994).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX   MEDLINE=95152076; PubMed=7849319;
RA   Gurney A.L., Kuang W.-J., Xie M.-H., Malloy B.E., Eaton D.L.,
RA   de Sauvage F.J.;
RT   "Genomic structure, chromosomal localization, and conserved
RT   alternative splice forms of thrombopoietin.";
RL   Blood 85:981-988(1995).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Liver;
RX   MEDLINE=96015174; PubMed=8537317;
RA   Kato T., Ogami K., Shimada Y., Iwamatsu A., Sohma Y., Akahori H.,
RA   Horie K., Kokubo A., Kudo Y., Maeda E., Kobayashi K., Ohashi H.,
RA   Ozawa T., Inoue H., Kawamura K., Miyazaki H.;
RT   "Purification and characterization of thrombopoietin.";
RL   J. Biochem. 118:229-236(1995).
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Placenta;
RX   MEDLINE=95122483; PubMed=7822271;
RA   Chang M., McNinch J., Basu R., Shutter J., Hsu R., Perkins C., Mar V.,
RA   Suggs S., Welcher A., Li L., Lu H., Bartley T., Hunt P., Martin F.,
RA   Samal B., Bogenberger J.;
RT   "Cloning and characterization of the human megakaryocyte growth and
RT   development factor (MGDF) gene.";
RL   J. Biol. Chem. 270:511-514(1995).
RN   [8]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RA   Im S.H., Lee W.S., Chung K.H.;
RT   "Cloning and sequencing of human thrombopoietin.";
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
CC       AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
CC       CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
CC       MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2/TPO-2 AND
```

```
CC       3/TRUNCATED; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
CC       TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
CC   -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC   -!- DATABASE: NAME=R&D Systems' cytokine source book: TPO;
CC       WWW="http://www.rndsystems.com/asp/g_sitebuilder.asp?bodyId=225".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L33410; AAA59857.1; -.
DR   EMBL; U11025; AAA50553.1; -.
DR   EMBL; L36051; AAC37568.1; -.
DR   EMBL; L36052; AAC37566.1; -.
DR   EMBL; D32046; BAA06807.1; -.
DR   EMBL; S76771; AAB33390.1; -.
DR   EMBL; D32047; BAA21930.1; -.
DR   EMBL; U59493; AAB03392.1; -.
DR   EMBL; U59494; AAB03393.1; -.
DR   EMBL; U59495; AAB03394.1; -.
DR   EMBL; U17071; AAA74083.1; -.
DR   PIR; S45331; S45331.
DR   Genew; HGNC:11795; THPO.
DR   MIM; 600044; -.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003978; thrombopoeitin.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR01485; THROMBOPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Cytokine; Glycoprotein; Hormone; Signal; Alternative splicing;
KW   Polymorphism.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    353       THROMBOPOIETIN.
FT   DISULFID    28    172       POTENTIAL.
FT   DISULFID    50    106       POTENTIAL.
FT   CARBOHYD   197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   206    206       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   255    255       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   340    340       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   348    348       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   133    136       MISSING (IN ISOFORM 2).
FT   VARSPLIC   160    198       MISSING (IN ISOFORM 3).
FT   VARIANT     14     14       L -> P (IN DBSNP:1042346).
FT                               /FTId=VAR_011795.
FT   VARIANT    116    116       G -> E (IN DBSNP:1126665).
FT                               /FTId=VAR_011796.
FT   CONFLICT    46     46       R -> K (IN REF. 8).
FT   CONFLICT    76     76       M -> MSQ (IN REF. 7).
FT   CONFLICT   113    113       Q -> E (IN REF. 2).
FT   CONFLICT   131    131       T -> P (IN REF. 7).
FT   CONFLICT   277    277       G -> E (IN REF. 8; AAB03393/AAB03394).
FT   CONFLICT   346    346       S -> C (IN REF. 8; AAB03393/AAB03394).
SQ   SEQUENCE   353 AA;  37822 MW;  F0AB5449B72E5526 CRC64;

  Query Match             10.5%;  Score 89;  DB 1;  Length 353;
  Best Local Similarity   26.3%;  Pred. No. 0.24;
  Matches   41;  Conservative   20;  Mismatches   75;  Indels   20;  Gaps    5;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              |||   || ||| : | ::    :  : | |   |   : :|    :   ||      :|
Db         24 APP--ACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPVLLPAVDFSLGEWKTQMEETKA 81

Qy         61 VEVWQGLALLSEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEA 118
               ::   : || | |:  |||      |     |   |   :  |:||      || |
Db         82 QDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLLLGALQSL------LGTQ--- 132

Qy        119 ISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKLK 154
                 ||       ||       :|  : : ||||::
Db        133 -LPPQG------RTTAHKDPNAIFLSFQHLLRGKVR 161


RESULT 13
POLG_HCVJ8
ID   POLG_HCVJ8     STANDARD;      PRT;  3033 AA.
AC   P26661;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
DE   Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
DE   (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
DE   (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
DE   (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
DE   NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
DE   NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
OS   Hepatitis C virus (isolate HC-J8) (HCV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11115;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92230232; PubMed=1314459;
RA   Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
RA   Fukuda S., Tsuda F., Mishiro S.;
RT   "Full-length sequence of a hepatitis C virus genome having poor
RT   homology to reported isolates: comparative study of four distinct
RT   genotypes.";
RL   Virology 188:331-341(1992).
CC   -!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
CC       HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC       NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC       precursor polyprotein, commonly with Asp or Glu in the P6
CC       position, Cys or Thr in P1 and Ser or Ala in P1'.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC       LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC       PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC       PROTEIN C AND MRNA.
CC   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10988; BAA01761.1; -.
DR   PIR; A40250; GNWVJ8.
DR   HSSP; P27958; 1HEI.
DR   MEROPS; S29.001; -.
DR   MEROPS; U39.001; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR002518; HCV_NS2.
DR   InterPro; IPR004109; HCV_NS3.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   Pfam; PF00998; HCV_RdRP; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
```

```
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   ProDom; PD186062; HCV_NS1; 1.
KW   Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
KW   Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
KW   Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
FT   INIT_MET      1      1       REMOVED FROM CAPSID PROTEIN C BY THE
FT                                CELLULAR AMINOPEPTIDASE.
FT   CHAIN         1    115       CAPSID PROTEIN C (POTENTIAL).
FT   CHAIN       116    191       MATRIX PROTEIN (POTENTIAL).
FT   CHAIN       192    383       MAJOR ENVELOPE PROTEIN E (POTENTIAL).
FT   CHAIN       384    733       NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
FT   CHAIN       734   1010       NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
FT   CHAIN      1011   1619       PROTEASE/HELICASE NS3 (POTENTIAL).
FT   CHAIN      1620   1866       NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
FT   CHAIN      1867   2017       NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
FT   CHAIN      2018   3033       RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
FT   TRANSMEM    347    369       POTENTIAL.
FT   ACT_SITE   1087   1087       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   1111   1111       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   1169   1169       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   NP_BIND    1234   1241       ATP (POTENTIAL).
FT   SITE       1320   1323       DECH BOX.
FT   CARBOHYD    196    196       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    209    209       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    233    233       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    299    299       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    305    305       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    417    417       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    423    423       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    430    430       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    448    448       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    477    477       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    534    534       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    542    542       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    558    558       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    578    578       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    627    627       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    649    649       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1091   1091       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2038   2038       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2359   2359       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2811   2811       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   3033 AA;  330177 MW;  1A173E7E3381FD1A CRC64;

  Query Match             9.5%;  Score 80.5;  DB 1;  Length 3033;
  Best Local Similarity   26.6%;  Pred. No. 24;
  Matches   41;  Conservative   22;  Mismatches   64;  Indels   27;  Gaps    7;

Qy          5 LICDSRVLERYLLEAKEAENITTGCAE---HCSLNENITV-PDTKVNFYAWKRM------ 54
              ::  | ||   :| |  |  : |||       ||: : | || :: : |:  |
Db       1659 IMTSSWVLAGGVLAAVAAYCLATGCISIIGRLHLNDRVVVAPDKEILYEAFDEMEECASK 1718

Qy         55 ----EVGQQAVEVWQG--LALLSEAVLRGQALLVNSSQPWEPLQ----LHVDKAVSGLRS 104
                  | ||:  |: :     || :|  : | :       |  |:     |:   :||::
Db       1719 AALIEEGQRMAEMLKSKIQGLLQQATRQAQDIQPAIQSSWPKLEQFWAKHMWNFISGIQY 1778

Qy        105 LTTLLRALG----AQKEAISPPDAASAAPLRTIT 134
              |  |    |    |   | |   ||  :|| | |
Db       1779 LAGLSTLPGNPAVASMMAFS---AALTSPLPTST 1809


RESULT 14
EP15_MOUSE
ID   EP15_MOUSE     STANDARD;      PRT;   897 AA.
AC   P42567;
DT   01-NOV-1995 (Rel. 32,.Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Epidermal growth factor receptor substrate 15 (Protein Eps15) (AF-1P
DE   protein).
GN   EPS15.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=93361014; PubMed=7689153;
RA   Fazioli F., Minichiello L., Matoskova B., Wong W.T., di Fiore P.P.;
RT   "eps15, a novel tyrosine kinase substrate, exhibits transforming
RT   activity.";
RL   Mol. Cell. Biol. 13:5814-5828(1993).
CC   -!- FUNCTION: INVOLVED IN CELL GROWTH REGULATION. MAY BE INVOLVED IN
CC       THE REGULATION OF MITOGENIC SIGNALS AND CONTROL OF CELL
CC       PROLIFERATION.
CC   -!- SUBUNIT: Interacts with Stonin 2.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- PTM: PHOSPHORYLATED ON TYROSINE BY EGFR.
CC   -!- SIMILARITY: CONTAINS 3 EH DOMAINS.
CC   -!- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L21768; AAA02912.1; -.
DR   HSSP; P02649; 1BZ4.
DR   MGD; MGI:104583; Eps15.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR000261; EPS15_repeat.
DR   InterPro; IPR003903; UIM.
DR   Pfam; PF00036; efhand; 4.
DR   SMART; SM00054; EFh; 3.
DR   SMART; SM00027; EH; 3.
DR   PROSITE; PS00018; EF_HAND; 2.
DR   PROSITE; PS50031; EH; 3.
KW   Phosphorylation; Calcium-binding; Repeat.
FT   DOMAIN       15    104       EH 1.
FT   DOMAIN      128    216       EH 2.
FT   CA_BIND     173    184       EF-HAND 1 (POTENTIAL).
FT   DOMAIN      224    314       EH 3.
FT   CA_BIND     236    247       EF-HAND 2 (POTENTIAL).
FT   DOMAIN      599    829       15 X 3 AA REPEATS OF D-P-F.
FT   REPEAT      599    601       1.
FT   REPEAT      623    625       2.
FT   REPEAT      629    631       3.
FT   REPEAT      634    636       4.
FT   REPEAT      640    642       5.
FT   REPEAT      645    647       6.
FT   REPEAT      651    653       7.
FT   REPEAT      665    667       8.
FT   REPEAT      673    675       9.
FT   REPEAT      693    695       10.
FT   REPEAT      711    713       11.
FT   REPEAT      806    808       12.
FT   REPEAT      827    829       13.
FT   MOD_RES     132    132       PHOSPHORYLATION (BY TYR-KINASES)
FT                                (POTENTIAL).
SQ   SEQUENCE   897 AA;  98470 MW;  08A0C0D423F873C2 CRC64;

  Query Match             9.3%;  Score 78.5;  DB 1;  Length 897;
  Best Local Similarity   26.5%;  Pred. No. 8.2;
  Matches   26;  Conservative   20;  Mismatches   35;  Indels   17;  Gaps    3;

Qy         12 LERYLLEAKEAENITTGCAEHC----SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGL 67
                || | | ::       |||      ||   ||  ::::: |  : ::  ::   : |
```

```
Db        408 LEEQLQEVRKK------CAEEAQLISSLKAEITSQESQISSYEEELLKAREELSRLQQET 461

Qy         68 ALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSL 105
              | | |:|  |:| |       ||| |: ::   : |:
Db        462 AQLEESVESGKAQL-------EPLQQHLQESQQEISSM 492


RESULT 15
ATCS_SYNP7
ID   ATCS_SYNP7     STANDARD;      PRT;   747 AA.
AC   P37279;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cation-transporting ATPase pacS (EC 3.6.3.-).
GN   PACS.
OS   Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=1140;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93380581; PubMed=8370468;
RA   Kanamaru K., Kashiwagi S., Mizuno T.;
RT   "The cyanobacterium, Synechococcus sp. PCC7942, possesses two
RT   distinct genes encoding cation-transporting P-type ATPases.";
RL   FEBS Lett. 330:99-104(1993).
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE OSMOTIC ADAPTATION.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC       (E1-E2 ATPASES). SUBFAMILY IB.
CC   -!- SIMILARITY: CONTAINS 1 HMA DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D16437; BAA03907.1; -.
DR   PIR; S36741; S36741.
DR   HSSP; Q04656; 1AW0.
DR   InterPro; IPR001757; ATPase_E1-E2.
DR   InterPro; IPR001934; HeavyMe_transpt.
DR   InterPro; IPR001454; Hlgnase/hydrlase.
DR   Pfam; PF00122; E1-E2_ATPase; 1.
DR   Pfam; PF00403; HMA; 1.
DR   Pfam; PF00702; Hydrolase; 1.
DR   PROSITE; PS00154; ATPASE_E1_E2; 1.
DR   PROSITE; PS01047; HMA_1; 1.
DR   PROSITE; PS50846; HMA_2; 1.
KW   Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
KW   Metal-binding.
FT   DOMAIN        1    101      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    102    122      POTENTIAL.
FT   DOMAIN      123    132      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    133    151      POTENTIAL.
FT   DOMAIN      152    158      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    159    179      POTENTIAL.
FT   DOMAIN      180    199      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    200    220      POTENTIAL.
FT   DOMAIN      221    348      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    349    371      POTENTIAL.
FT   DOMAIN      372    378      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    379    396      POTENTIAL.
FT   DOMAIN      397    688      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    689    708      POTENTIAL.
FT   DOMAIN      709    720      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    721    739      POTENTIAL.
FT   DOMAIN      740    747      CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        4     70      HMA.
FT   MOD_RES     434    434      PHOSPHORYLATION (BY SIMILARITY).
FT   METAL        14     14      POTENTIAL.
FT   METAL        17     17      POTENTIAL.
FT   METAL       634    634      MAGNESIUM (BY SIMILARITY).
FT   METAL       638    638      MAGNESIUM (BY SIMILARITY).
SQ   SEQUENCE   747 AA;  79732 MW;  0AF148B92EB85CEE CRC64;

  Query Match             9.2%;  Score 78;  DB 1;  Length 747;
  Best Local Similarity   25.4%;  Pred. No. 7.3;
  Matches   33;  Conservative   22;  Mismatches   45;  Indels   30;  Gaps    5;

Qy         17 LEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQAV---EVWQGLALLSEA 73
              |:|| |  :  |          : :|:| |:|    | |:  |::     ||  | : : |:
Db        233 LQAKTARVLRQG--------QELTLPITEVQVEDWVRVRPGEKVPVDGEVIDGRSTVDES 284

Qy         74 VLRGQALLVNSSQPWEPLQLHVDKAVSGLR-----SLTTLLRALG-----AQKEAISPPD 123
              :: |::|         |:|  |   | |       |||     :|     ||   :
Db        285 MVTGESL---------PVQKQVGDEVIGATLNKTGSLTIRATRVGRETFLAQIVQLVQQA 335

Qy        124 AASAAPLRTI 133
               || ||:: :
Db        336 QASKAPIQRL 345


Search completed: June 13, 2003, 15:04:03
Job time : 14 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:04:06 ; Search time 34 Seconds
(without alignments)
999.935 Million cell updates/sec

Title: US-09-830-964-1
Perfect score: 846
Sequence: 1 APPRLICDSRVLERYLLEAK.....SNFLRGKLKLYTGEACRTGD 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	680.5	80.4	195	6	Q9GKA3	Q9gka3 oryctolagus
2	680.5	80.4	195	6	Q9GKA2	Q9gka2 oryctolagus
3	678	80.1	194	6	Q9MYM8	Q9mym8 sus scrofa
4	188	22.2	50	11	Q9QV40	Q9qv40 rattus sp.
5	88	10.4	323	16	Q8ZDC8	Q8zdc8 yersinia pe
6	87.5	10.3	346	16	Q8ZKZ4	Q8zkz4 salmonella
7	87.5	10.3	346	16	Q8Z2M5	Q8z2m5 salmonella
8	85	10.0	3722	2	P94873	P94873 lysobacter
9	83	9.8	296	16	Q8ZAY4	Q8zay4 yersinia pe
10	83	9.8	339	16	Q9HZM7	Q9hzm7 pseudomonas
11	82.5	9.8	3033	12	Q9DHD6	Q9dhd6 hepatitis c
12	82	9.7	815	10	Q9FK91	Q9fk91 arabidopsis
13	79.5	9.4	1829	16	O86808	O86808 streptomyce
14	79.5	9.4	3033	12	Q9IZA1	Q9iza1 hepatitis c
15	79	9.3	480	10	Q43380	Q43380 avena sativ
16	79	9.3	2364	5	Q22896	Q22896 caenorhabdi
17	78.5	9.3	813	16	Q8Z9B9	Q8z9b9 salmonella
18	78.5	9.3	3019	12	O92529	O92529 hepatitis c
19	78	9.2	348	16	O86751	O86751 streptomyce
20	78	9.2	378	17	Q8ZZK3	Q8zzk3 pyrobaculum
21	78	9.2	455	16	Q8UBQ9	Q8ubq9 agrobacteri
22	78	9.2	548	2	Q93T48	Q93t48 buchnera ap
23	78	9.2	548	2	Q93N35	Q93n35 buchnera ap
24	77.5	9.2	242	16	Q8YY76	Q8yy76 anabaena sp
25	77.5	9.2	310	16	Q9L033	Q91033 streptomyce
26	77.5	9.2	451	16	Q8XSE8	Q8xse8 ralstonia s
27	77.5	9.2	1003	10	Q9FIR4	Q9fir4 arabidopsis
28	77.5	9.2	2556	11	Q91XW2	Q91xw2 mus musculu
29	77	9.1	379	2	Q9L9F5	Q919f5 streptomyce
30	77	9.1	383	16	Q92QV7	Q92qv7 rhizobium m
31	77	9.1	447	2	P72270	P72270 rhodococcus
32	77	9.1	451	16	P74054	P74054 synechocyst
33	76.5	9.0	154	16	Q9PGA7	Q9pga7 xylella fas
34	76.5	9.0	425	16	Q8YF20	Q8yf20 brucella me
35	76.5	9.0	619	3	Q96VB7	Q96vb7 botrytis ci
36	76.5	9.0	1315	3	Q8X215	Q8x215 botrytis ci
37	76.5	9.0	1315	3	Q8X1E7	Q8x1e7 botrytis ci
38	75.5	8.9	637	16	Q55413	Q55413 synechocyst
39	75.5	8.9	871	5	Q9NB04	Q9nb04 drosophila
40	75	8.9	308	10	Q94I41	Q94i41 oryza sativ
41	75	8.9	397	8	Q9TNG5	Q9tng5 coriaria te
42	75	8.9	548	2	Q93N34	Q93n34 buchnera ap
43	75	8.9	554	2	Q9RPH5	Q9rph5 mycobacteri
44	74.5	8.8	361	6	Q9TSW6	Q9tsw6 sus scrofa
45	74.5	8.8	389	2	Q9AN18	Q9an18 bradyrhizob

ALIGNMENTS

RESULT 1
Q9GKA3
ID Q9GKA3 PRELIMINARY; PRT; 195 AA.
AC Q9GKA3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Erythropoietin.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21290682; PubMed=11396976;
RA Vilalta A., Wu D., Margalith M., Hobart P.;
RT "Rabbit EPO Gene and cDNA: Expression of Rabbit EPO after
RT Intramuscular Injection of pDNA.";
RL Biochem. Biophys. Res. Commun. 284:823-827(2001).
DR EMBL; AF290943; AAG36961.1; -.
DR HSSP; P01588; 1CN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythroptn.
DR Pfam; PF00758; EPO_TPO; 1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO; 1.
SQ SEQUENCE 195 AA; 21053 MW; 0999DA7D852713F3 CRC64;

Query Match 80.4%; Score 680.5; DB 6; Length 195;
Best Local Similarity 81.3%; Pred. No. 1.4e-60;
Matches 135; Conservative 12; Mismatches 18; Indels 1; Gaps 1;

```
Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              || |||||||||||:||||||||:| |||| ||| |||||||||||: ||: | |: |
Db         29 APARLICDSRVLERYILEAKEAENVTMGCAEGCSLGENITVPDTKVNFHHWKKSEAGRHA 88

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              |||||||||||||:|| |||| ||||  | ||:|||||||||||||:|||||| ||||:|
Db         89 VEVWQGLALLSEAMLRSQALLANSSQLPETLQVHVDKAVSGLRSLTSLLRALGVQKEAVS 148
```

```
Qy        121 PPDAA-SAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              ||:|| ||||||| ||| ||||:||||||||||||||||| ||
Db        149 PPEAASSAAPLRTVAADTLCKLFRIYSNFLRGKLKLYTGEACRRGD 194


RESULT 2
Q9GKA2
ID   Q9GKA2      PRELIMINARY;      PRT;   195 AA.
AC   Q9GKA2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Erythropoietin.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21290682; PubMed=11396976;
RA   Vilalta A., Wu D., Margalith M., Hobart P.;
RT   "Rabbit EPO Gene and cDNA: Expression of Rabbit EPO after
RT   Intramuscular Injection of pDNA.";
RL   Biochem. Biophys. Res. Commun. 284:823-827(2001).
DR   EMBL; AF290944; AAG36962.1; -.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
SQ   SEQUENCE   195 AA;  21025 MW;  1F1DC7F403A303EC CRC64;

  Query Match             80.4%;  Score 680.5;  DB 6;  Length 195;
  Best Local Similarity   81.3%;  Pred. No. 1.4e-60;
  Matches  135;  Conservative   12;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              || |||||||||||||:||||||||:| |||| ||| ||||||||||||: ||: | |: |
Db         29 APARLICDSRVLERYILEAKEAENVTMGCAEGCSLGENITVPDTKVNFHHWKKSEAGRHA 88

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              |||||||||||||:|| |||| |||| | ||:|||||||||||||:|||||| |||||:|
Db         89 VEVWQGLALLSEAMLRSQALLANSSQLPETLQVHVDKAVSGLRSLTSLLRALGVQKEAVS 148

Qy        121 PPDAA-SAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
              ||:|| ||||||| ||| ||||:||||||||||||||||| ||
Db        149 PPEAASSAAPLRTVAADTLCKLFRIYSNFLRGKLKLYTGEACRRGD 194


RESULT 3
Q9MYM8
ID   Q9MYM8      PRELIMINARY;      PRT;   194 AA.
AC   Q9MYM8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Erythropoietin precursor.
GN   EPO.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NORWEGIAN LANDRACE; TISSUE=KIDNEY;
RA   David B., Harbitz I.;
RT   "The porcine erythropoietin gene: cDNA and genomic sequences and
RT   expression analyses.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ249745; CAB96416.1; -.
DR   EMBL; AJ249746; CAB96417.1; -.
DR   HSSP; P01588; 1CN4.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
DR   PROSITE; PS00817; EPO_TPO; 1.
KW   Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    194       ERYTHROPOIETIN.
SQ   SEQUENCE   194 AA;  21303 MW;  77881A6F6F20EA1C CRC64;

  Query Match             80.1%;  Score 678;  DB 6;  Length 194;
  Best Local Similarity   82.0%;  Pred. No. 2.5e-60;
  Matches  137;  Conservative    7;  Mismatches   21;  Indels    2;  Gaps    1;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQA 60
              ||||||||||||||||:||||| || | |||| || :|||||||||||||||||||| |||
Db         27 APPRLICDSRVLERYILEAKEGENATMGCAESCSFSENITVPDTKVNFYAWKRMEVQQQA 86

Qy         61 VEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAIS 120
              :|||||||||||||:|:|||||| ||||| | ||||||||||||||||:||||||||||||
Db         87 MEVWQGLALLSEAILQGQALLANSSQPSEALQLHVDKAVSGLRSLTSLLRALGAQKEAIP 146

Qy        121 PPDA--ASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEACRTGD 165
               |||  :|| ||||   ||  |||| ||||||||| ||||||||  |
Db        147 LPDASPSSATPLRTFAVDTLCKLFRNYSNFLRGKLTLYTGEACRRRD 193


RESULT 4
Q9QV40
ID   Q9QV40      PRELIMINARY;      PRT;    50 AA.
AC   Q9QV40;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Erythropoietin (Fragment).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=94115047; PubMed=7764337;
RA   Okano M., Suga H., Masuda S., Nagao M., Narita H., Ikura K.,
RA   Sasaki R.;
RL   Biosci. Biotechnol. Biochem. 57:1882-1885(1993).
DR   HSSP; P01588; 1EER.
DR   InterPro; IPR001323; EPO_TPO.
DR   InterPro; IPR003013; Erythroptn.
DR   Pfam; PF00758; EPO_TPO; 1.
DR   PRINTS; PR00272; ERYTHROPTN.
SQ   SEQUENCE   50 AA;  5587 MW;  70B44A8BFE016034 CRC64;

  Query Match             22.2%;  Score 188;  DB 11;  Length 50;
  Best Local Similarity   78.0%;  Pred. No. 9.9e-12;
  Matches   39;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYA 50
              |||||| ||||||||:||||||| :| | ||   |:| ||||||||||||
Db          1 APPRLIXDSRVLERYILEAKEAEXVTMGXAEGPRLSEXITVPDTKVNFYA 50


RESULT 5
Q8ZDC8
ID   Q8ZDC8      PRELIMINARY;      PRT;   323 AA.
AC   Q8ZDC8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Ribonucleoside-diphosphate reductase 2 beta chain (EC 1.17.4.1).
GN   NRDF OR YPO2648.
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
```

```
OC   Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX   MEDLINE=21470413; PubMed=11586360;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL   Nature 413:523-527(2001).
DR   EMBL; AJ414153; CAC92889.1; -.
DR   InterPro; IPR000358; Ribonucl_redctse.
DR   Pfam; PF00268; ribonuc_red_sm; 1.
DR   PROSITE; PS00368; RIBORED_SMALL; 1.
KW   Oxidoreductase; Complete proteome.
SQ   SEQUENCE   323 AA;  36823 MW;  87C21F7BBB9B7FD2 CRC64;

  Query Match             10.4%;  Score 88;  DB 16;  Length 323;
  Best Local Similarity   25.2%;  Pred. No. 1.3;
  Matches   34;  Conservative   20;  Mismatches   59;  Indels   22;  Gaps    5;

Qy         38 NITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVD- 96
              |:  | |:::   | ::|   : :|||    | |   |  :  | |    |  |  |
Db          2 NVVKPITRISAINWNKIE-DDKDLEVWN--RLTSNFWLPEKVPLSNDIPSWATLTPHEQQ 58

Qy         97 ---KAVSGLRSLTTLLRALGAQ---KEAISPPDAASAAPLRTITADTFRKLFRVYSNFLR 150
                 :  :||  | |:   |||    |:||:| : |  : :  : |   |    ::|
Db         59 LTIRVFTGLTLLDTIQNTLGAPALIKDAITPHEEAIFSNISFMEAVHARSYSSIFSTL-- 116

Qy        151 GKLKLYTGEACRTGD 165
                       | | |
Db        117 ----------CLTSD 121


RESULT 6
Q8ZKZ4
ID   Q8ZKZ4      PRELIMINARY;      PRT;   346 AA.
AC   Q8ZKZ4;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Periplasmic sensor in multi-comopnent regulatory system with TorS
DE   (sensory kinase) and TorR (regulator), regulates tor operon.
GN   TORT OR STM3825.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
DR   EMBL; AE008878; AAL22684.1; -.
DR   InterPro; IPR001761; PeriplaBP/Lac1.
DR   Pfam; PF00532; Peripla_BP_like; 1.
KW   Complete proteome.
SQ   SEQUENCE   346 AA;  38440 MW;  E37CAB58E49FD716 CRC64;

  Query Match             10.3%;  Score 87.5;  DB 16;  Length 346;
  Best Local Similarity   26.7%;  Pred. No. 1.5;
  Matches   44;  Conservative   22;  Mismatches   48;  Indels   51;  Gaps    9;

Qy         10 RVLERYLLEAKEAENITTG--CAEHCSLNE--NITVPDTKVNFYAWKRMEVGQQAVEVWQ 65
              | | : :||     |:  |   |   :: |  |:| | | |:||          :|::
Db        217 RNLLQEMLERHPDANVVAGSAIAAEAAMGEGRNLTTPLTIVSFYL---------THQVYR 267

Qy         66 GLALLSEAVLRGQALLVNSSQ-PWEPLQLHVDKAVSGLRSLTTLLRALGAQ--KEAISPP 122
              ||        ||  |:  | |  |:           |  ::|  :: |  |   | ||||
Db        268 GLK-------RGHILMALSDQMAWQ----------GELAITQSIKVLQGQPVPENISPP 309

Qy        123 -------DAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEA 160
                    :| ||   |:::   || ::          |  || ||
Db        310 VLILTHNNADSARVRRSLSPPGFRPVY---------LYQYTSEA 344


RESULT 7
Q8Z2M5
ID   Q8Z2M5      PRELIMINARY;      PRT;   346 AA.
AC   Q8Z2M5;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Solute binding receptor protein.
GN   STY3952.
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
DR   EMBL; AL627280; CAD03169.1; -.
DR   InterPro; IPR001761; PeriplaBP/Lac1.
DR   Pfam; PF00532; Peripla_BP_like; 1.
KW   Receptor; Complete proteome.
SQ   SEQUENCE   346 AA;  38546 MW;  F80FB1688BC83A8F CRC64;

  Query Match             10.3%;  Score 87.5;  DB 16;  Length 346;
  Best Local Similarity   26.7%;  Pred. No. 1.5;
  Matches   44;  Conservative   22;  Mismatches   48;  Indels   51;  Gaps    9;

Qy         10 RVLERYLLEAKEAENITTG--CAEHCSLNE--NITVPDTKVNFYAWKRMEVGQQAVEVWQ 65
              | | : :||     |:  |   |   :: |  |:| | | |:||          :|::
Db        217 RNLLQEMLERHPDANVVAGSAIAAEAAMGEGRNLTTPLTIVSFYL---------THQVYR 267

Qy         66 GLALLSEAVLRGQALLVNSSQ-PWEPLQLHVDKAVSGLRSLTTLLRALGAQ--KEAISPP 122
              ||        ||  |:  | |  |:           |  ::|  :: |  |   | ||||
Db        268 GLK-------RGHILMALSDQMAWQ----------GELAITQSIKVLQGQPVPENISPP 309

Qy        123 -------DAASAAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEA 160
                    :| ||   |:::   || ::          |  || ||
Db        310 VLILTHNNADSARVRRSLSPPGFRPVY---------LYQYTSEA 344


RESULT 8
P94873
ID   P94873      PRELIMINARY;      PRT;  3722 AA.
AC   P94873;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
```

```
DE   Alpha-aminoadipyl-cysteinyl-valine synthetase.
GN   PCBAB.
OS   Lysobacter lactamgenus.
OC   Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC   Lysobacter.
OX   NCBI_TaxID=39596;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YK90;
RX   MEDLINE=96275949; PubMed=8737573;
RA   Kimura H., Miyashita H., Sumino Y.;
RT   "Organization and expression in Pseudomonas putida of the gene cluster
RT   involved in cephalosporin biosynthesis from Lysobacter lactamgenus
RT   YK90.";
RL   Appl. Microbiol. Biotechnol. 45:490-501(1996).
DR   EMBL; D50308; BAA08846.1; -.
DR   HSSP; P14687; 1AMU.
DR   InterPro; IPR002106; AAtRNA_ligaseII.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR001242; Condensatn.
DR   InterPro; IPR000977; DNA_ligase.
DR   InterPro; IPR003880; Ppantne_attach.
DR   InterPro; IPR000379; Ser_estrs_site.
DR   InterPro; IPR001031; Thioesterase.
DR   Pfam; PF00501; AMP-binding; 3.
DR   Pfam; PF00668; Condensation; 3.
DR   Pfam; PF00550; pp-binding; 3.
DR   Pfam; PF00975; Thioesterase; 1.
DR   PROSITE; PS00179; AA_TRNA_LIGASE_II_1; UNKNOWN_1.
DR   PROSITE; PS50075; ACP_DOMAIN; 3.
DR   PROSITE; PS00455; AMP_BINDING; 1.
DR   PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_2.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_2.
KW   Phosphopantetheine.
SQ   SEQUENCE   3722 AA;  411607 MW;  3597B3483463809B CRC64;

  Query Match             10.0%;  Score 85;  DB 2;  Length 3722;
  Best Local Similarity   22.1%;  Pred. No. 57;
  Matches   34;  Conservative   18;  Mismatches   50;  Indels   52;  Gaps    5;

Qy         16 LLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQAVEV----WQGLALLS 71
              |||   |: :              :| |  :|  |:| ::  | ::     ||
Db       2138 LLEGATAQTLP-------------PLPATYADFSVWQRQQLSDQRLDALFDYWQ------ 2178

Qy         72 EAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQ-----KEAISPPDAAS 126
                        |   |:|||| :|              |  ||     :| :   || :
Db       2179 -----------RSLAGWQPLQLPLD-------------HARPAQFDYLGREIVFDVDATT 2214

Qy        127 AAPLRTITADTFRKLFRVYSNFLRGKLKLYTGEA 160
                 || :   |    | |        || |:|::
Db       2215 CDQLRVLAQTTRTSFFSVLLAAYYLTLKAYSGQS 2248


RESULT 9
Q8ZAY4
ID   Q8ZAY4      PRELIMINARY;      PRT;   296 AA.
AC   Q8ZAY4;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative 2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60).
GN   YPO3648.
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX   MEDLINE=21470413; PubMed=11586360;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL   Nature 413:523-527(2001).
DR   EMBL; AJ414158; CAC93117.1; -.
DR   InterPro; IPR002204; 3hydroxisobut_dh.
DR   InterPro; IPR001744; 6PGD.
DR   Pfam; PF03446; NAD_binding_2; 1.
DR   PROSITE; PS00895; 3_HYDROXYISOBUT_DH; UNKNOWN_1.
KW   Oxidoreductase; Hypothetical protein; Complete proteome.
SQ   SEQUENCE   296 AA;  32140 MW;  C883AFCC6868429D CRC64;

  Query Match             9.8%;  Score 83;  DB 16;  Length 296;
  Best Local Similarity   24.2%;  Pred. No. 3.6;
  Matches   48;  Conservative   34;  Mismatches   62;  Indels   54;  Gaps   10;

Qy          6 ICDSRVLERYLLEAK-EAENITTGCAEHCSLNENIT-VPDTKVNFYAWKRMEVGQQAVEV 63
              :||: : :  :| |:  ||::|   |   ||:  || :|:::    | ::: :|   :
Db         32 LCDADITKVQILTAELNAESLTVDNA--ASLDALITMLPNSE----AVEQVLLGSDGISG 85

Qy         64 WQGLALLSEAVLRGQALLVNSSQP-----------WE------PLQLHVDKAVSG---- 101
              |  :| ||:| :    : ::|| |            ||      |:   | || :|
Db         86 W--VAQLSQAAV---VIDMSSSDPERSRRLAILLAVWELDYLDAPVSGGVKKAQNGTLSI 140

Qy        102 --------LRSLTTLLRALGAQKEAISPPDAASAAP----------LRTITADTFRKLF 142
                    |:|  | | |:| |   : |  :  ||              | || |    : |
Db        141 LIGGEDRVLKSCYTALAAMGEQILFVGPAGSGHAAKALNNYVSATGLLATIEALHVAQRF 200

Qy        143 RVYSNFLRGKLKLYTGEA 160
                :    :   |   || :
Db        201 GIEPEVMTEVLNTSTGRS 218


RESULT 10
Q9HZM7
ID   Q9HZM7      PRELIMINARY;      PRT;   339 AA.
AC   Q9HZM7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   UDP-N-acetylpyruvoylglucosamine reductase.
GN   MURB OR PA2977.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004723; AAG06365.1; -.
DR   HSSP; P08373; 2MBR.
DR   InterPro; IPR003170; MurB.
DR   InterPro; IPR001575; Oxid_FAD_bind.
DR   InterPro; IPR000531; TonB_boxC.
DR   Pfam; PF01565; FAD_binding_4; 1.
DR   Pfam; PF02873; MurB_C; 1.
DR   PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   339 AA;  37627 MW;  C0C8EF9F2938FE27 CRC64;

  Query Match             9.8%;  Score 83;  DB 16;  Length 339;
```

```
  Best Local Similarity   22.7%;  Pred. No. 4.3;
  Matches   41;  Conservative   22;  Mismatches   48;  Indels   70;  Gaps    8;

Qy         31 EHCSLNE-NITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLV------- 82
              |||||   |    | :    |  | |      :| : |||  |   ||  |||
Db          7 EHCSLKPYNTFGIDVRARLLAHARDE-----ADVREALALARE---RGLPLLVIGGGSNL 58

Qy         83 --------------------------------NSSQPWEP-LQLHVDKAVSGLRSLTTLL 109
                                              : : |:| :|  ::: ::|| :|: :
Db         59 LLTRDVEALVLRMASQGRRIVSDAADSVLVEAEAGEAWDPFVQWSLERGLAGLENLSLI- 117

Qy        110 RALGAQKEAISPPDAASAAPLRTITA------DTFRKLFRVYSNFLRGKLKLYTGEACRT 163
                    |   |||:: | |      | |  |  :   :    | |: :  :|||
Db        118 ------------PGTVGAAPMQNIGAYGVELKDVFDSLTAL--DRQDGTLREFDRQACRF 163

Qy        164 G 164
              |
Db        164 G 164


RESULT 11
Q9DHD6
ID   Q9DHD6      PRELIMINARY;      PRT;  3033 AA.
AC   Q9DHD6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE   (NS1)].
OS   Hepatitis C virus type 2b.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=31650;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA;
RX   MEDLINE=21296595; PubMed=11402859;
RA   Murakami K., Abe M., Kageyama T., Kamoshita N., Nomoto A.;
RT   "Down-regulation of translation driven by hepatitis C virus internal
RT   ribosomal entry site by the 3' untranslated region of RNA.";
RL   Arch. Virol. 146:729-741(2001).
CC   -!- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR   EMBL; AB030907; BAB08107.1; -.
DR   HSSP; P27958; 1A1V.
DR   InterPro; IPR000345; CytC_heme_bind.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR002518; HCV_NS2.
DR   InterPro; IPR004109; HCV_NS3.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; HCV_RdRP; 1.
DR   ProDom; PD186062; HCV_NS1; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; RNA-directed RNA polymerase; Transmembrane.
FT   CHAIN         1    191       CORE PROTEIN.
FT   CHAIN       192    383       E1 PROTEIN.
FT   CHAIN       384    750       E2 PROTEIN.
FT   CHAIN       751    813       P7 PROTEIN.
FT   CHAIN       814   1030       NS2 PROTEIN.
FT   CHAIN      1031   1661       NS3 PROTEINASE/HELICASE.
FT   CHAIN      1662   1715       NS4A PROTEIN.
FT   CHAIN      1716   1976       NS4B PROTEIN.
FT   CHAIN      1977   2442       NS5A PROTEIN.
FT   CHAIN .    2443   3033       NS5B RNA-DEPENDENT RNA POLYMERASE.
SQ   SEQUENCE   3033 AA;  329981 MW;  6B183FED090872B4 CRC64;

  Query Match             9.8%;  Score 82.5;  DB 12;  Length 3033;
  Best Local Similarity   27.0%;  Pred. No. 78;
  Matches   40;  Conservative   21;  Mismatches   60;  Indels   27;  Gaps    7;

Qy         11 VLERYLLEAKEAENITTGCAE---HCSLNENITV-PDTKVNFYAWKRM----------EV 56
              ||   :| |  |  : |||        ||: : | || :: : |:  |           |
Db       1665 VLAGGVLAAVAAYCLATGCISIIGRIHLNDQVVVAPDKEILYEAFDEMEECASKAALIEE 1724

Qy         57 GQQAVEVWQG--LALLSEAVLRGQALLVNSSQPWEPLQ----LHVDKAVSGLRSLTTLLR 110
              ||:  |: :   | || :|  : | :        |  ::     |:   :||:: |  |
Db       1725 GQRMAEMLKSKILGLLQQATKQAQDIQPAMQSSWPKIEQFWARHMWNFISGIQYLAGLST 1784

Qy        111 ALG----AQKEAISPPDAASAAPLRTIT 134
                |    |   | |   ||  :|| | |
Db       1785 LPGNPAVASMMAFS---AALTSPLPTST 1809


RESULT 12
Q9FK91
ID   Q9FK91      PRELIMINARY;      PRT;  815 AA.
AC   Q9FK91;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Seed maturation protein PM38 protein.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   MEDLINE=98403884; PubMed=9734815;
RA   Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT   Sequence features of the regions of 1,367,185 bp covered by 19
RT   physically assigned P1 and TAC clones.";
RL   DNA Res. 5:203-216(1998).
DR   EMBL; AB012244; BAB09119.1; -.
DR   InterPro; IPR002086; Aldehyde_dehydr.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE   815 AA;  91534 MW;  4B1B602057D46B60 CRC64;

  Query Match             9.7%;  Score 82;  DB 10;  Length 815;
  Best Local Similarity   21.8%;  Pred. No. 16;
  Matches   42;  Conservative   29;  Mismatches   52;  Indels   70;  Gaps   12;

Qy         21 EAENITTGCA----------EHCSLNENITVPDTKVNFYAWKRMEVG-----QQAVEVW 64
              | : :   ||          |:| : : :||||: :| | :|| :||     :: :| |
Db        328 EKDGLLYNCAFSICDLGKGRNEYC-IMQLVTVPDSNLNMY-FKRGKVGDDPNAEERLEEW 385

Qy         65 QGLALLSEAVLRGQALLV-----NSSQPWE------------PLQ----LHVDKAVSGL 102
              :     || ::  | |       | :|||                |:     : |    ||
Db        386 ED----EEAAIKEFARLFEEIAGNEFEPWEREKKIQKKPHKFFPIDMDDGIEVRSGALGL 441
```

```
Qy        103 RSL-------------TTLLRALGAQ--------KEAISPPDAASAAPLRTITADTFRKL 141
              | |                 :: |  |        :  : |||     |:  :|    ::
Db        442 RQLGIASAHCKLDSFVANFIKVLCGQEIYNYALMELGLDPPD----LPMGMLTDIHLKRC 497

Qy        142 FRVYSNFLRGKLK 154
                  |   |:  |:|
Db        498 EEVLLEFVE-KVK 509


RESULT 13
O86808
ID   O86808      PRELIMINARY;      PRT;  1829 AA.
AC   O86808;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative sensory histidine kinase.
GN   SCO5748 OR SC7C7.03.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Harris D., Taylor K.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
CC   -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC       KINASES.
DR   EMBL; AL031031; CAA19849.1; -.
DR   HSSP; P52934; 1DZ3.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR004358; Bact_sens_pr_C.
DR   InterPro; IPR003018; GAF.
DR   InterPro; IPR003660; HAMP.
DR   InterPro; IPR003661; His_kinA.
DR   InterPro; IPR004359; HIS_KIN_sig.
DR   InterPro; IPR001789; Response_reg.
DR   Pfam; PF01590; GAF; 1.
DR   Pfam; PF00672; HAMP; 11.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00072; response_reg; 1.
DR   Pfam; PF00512; signal; 1.
DR   PRINTS; PR00344; BCTRLSENSOR.
DR   ProDom; PD000039; Response_reg; 1.
DR   SMART; SM00065; GAF; 1.
DR   SMART; SM00304; HAMP; 12.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00388; HisKA; 1.
DR   SMART; SM00448; REC; 1.
KW   Kinase; Phosphorylation; Sensory transduction; Transferase.
SQ   SEQUENCE   1829 AA;  195751 MW;  C9C8699938C956A4 CRC64;

  Query Match             9.4%;  Score 79.5;  DB 16;  Length 1829;
  Best Local Similarity   22.8%;  Pred. No. 82;
  Matches   37;  Conservative   29;  Mismatches   61;  Indels   35;  Gaps    6;

Qy         18 EAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALL------- 70
              : :|  ::||  | :  |:: :|||         | || | |  : |    |
Db        328 QVREISHVTTAVA-NGDLSKKVTVP---------ARGEVAQLAETINQMTETLRIFADEV 377

Qy         71 --------SEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAISPP 122
                      |  | ||| :  ::  |: |    |   :  |:||| :| : |   |::
Db        378 TRVANETGGEGQLGGQANVPGAAGIWKDL---TDSVNTVFRNLTTQVRDIAAVTTAVASG 434

Qy        123 DAASAAPLRTITADTFRKLFRVYS--NFLRGKLKLYTGEACR 162
              | :     : :| |   ::  : :  | :  :|  :  |  |
Db        435 DLS-----QKVTVDVAGEMLELKNTVNTMVDQLSAFGAEVTR 471


RESULT 14
Q9IZA1
ID   Q9IZA1      PRELIMINARY;      PRT;  3033 AA.
AC   Q9IZA1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE   (NS1)].
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD2B-1;
RA   Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
RA   Sato C.;
RT   "Full length cDNA sequence of HCV genotype 2b, strain MD2b-1.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR   EMBL; AF238486; AAF59945.1; -.
DR   HSSP; P27958; 1HEI.
DR   MEROPS; S29.001; -.
DR   MEROPS; U39.001; -.
DR   InterPro; IPR000345; CytC_heme_bind.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR002518; HCV_NS2.
DR   InterPro; IPR004109; HCV_NS3.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; HCV_RdRP; 1.
```

```
DR   ProDom; PD186062; HCV_NS1; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; RNA-directed RNA polymerase; Transmembrane.
SQ   SEQUENCE   3033 AA;  330723 MW;  FE04FEC7C385A13A CRC64;

  Query Match             9.4%;  Score 79.5;  DB 12;  Length 3033;
  Best Local Similarity   26.6%;  Pred. No. 1.6e+02;
  Matches   41;  Conservative   22;  Mismatches   64;  Indels   27;  Gaps    7;

Qy          5 LICDSRVLERYLLEAKEAENITTGCAE---HCSLNENITV-PDTKVNFYAWKRM------ 54
              ::  | ||   :| |  |  : |||        ||: : | || :: : |:  |
Db       1659 IMTSSWVLAGGVLAAVAAYCLATGCISIIGRLHLNDQVVVAPDKEILYEAFDEMEECASK 1718

Qy         55 ----EVGQQAVEVWQG--LALLSEAVLRGQALLVNSSQPWEPLQ----LHVDKAVSGLRS 104
                  | ||:  |: :     || :|  : | :       |  |:     |:   :||::
Db       1719 AALIEEGQRMAEMLKSKIQGLLQQATRQAQDIHPAIQSSWPKLEQFWAKHMWNFISGIQY 1778

Qy        105 LTTLLRALG----AQKEAISPPDAASAAPLRTIT 134
              |  |    |    |   | |   ||  :|| | |
Db       1779 LAGLSTLPGNPAVASMMAFS---AALTSPLPTST 1809


RESULT 15
Q43380
ID   Q43380      PRELIMINARY;      PRT;   480 AA.
AC   Q43380;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Putative pp70 ribosomal protein S6 kinase.
GN   ASPK11.
OS   Avena sativa (Oat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Aveneae; Avena.
OX   NCBI_TaxID=4498;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. RHIANNON; TISSUE=ALEURONE;
RX   MEDLINE=95284341; PubMed=7766874;
RA   Huttly A.K., Philips A.L.;
RT   "gibberellin regulated expression in oat aleurone cells of two kinases
RT   that hsow homology to nap kinase and a ribosomal protein kinase.";
RL   Plant Mol. Biol. 27:1043-1052(1995).
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR   EMBL; X79992; CAA56313.1; -.
DR   HSSP; P05132; 1CTP.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR000961; Pkinase_C.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00433; pkinase_C; 1.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00133; S_TK_X; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Ribosomal protein;
KW   Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   480 AA;  53532 MW;  33596A101DBB077D CRC64;

  Query Match             9.3%;  Score 79;  DB 10;  Length 480;
  Best Local Similarity   22.4%;  Pred. No. 17;
  Matches   32;  Conservative   26;  Mismatches   65;  Indels   20;  Gaps    4;

Qy         30 AEHCSLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEA--VLRGQALLVNSSQP 87
              | |     :|  ||:           : :   ||:    ::::    | |  |||: | |
Db         51 AVHTPTEASIVTPDSPA--------PLAESGEEVYNDPVVITKRSHSLVGPTLLVSQSLP 102

Qy         88 WEPLQLHVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKL------ 141
                | ||  :: | |  |  | :   : :||:|  : ::      :  |  |  |

Db        103 LSKLTLHESESSSDL--LECLSKEKQSNQEALSDEELSNTKENEAVGLDNFEVLKLVGQG 160

Qy        142 --FRVYSNFLRGKLKLYTGEACR 162
                 :||   ::|  ::|  :  |
Db        161 AFGKVYQVRMKGTSEIYAMKVMR 183


Search completed: June 13, 2003, 15:13:05
Job time : 43 secs
```

STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact ***the searcher or contact:***

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

- *I am an examiner in Workgroup:* ☐ *Example: 1610*
- *Relevant prior art **found**, search results used as follows:*
  - ☐ 102 rejection
  - ☐ 103 rejection
  - ☐ Cited as being of interest.
  - ☐ Helped examiner better understand the invention.
  - ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
  (journal articles, conference proceedings, new product announcements etc.)

- *Relevant prior art **not found:***
  - ☐ Results verified the lack of relevant prior art (helped determine patentability).
  - ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 – Circ. Desk

STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 96531

TO:	From:
Patricia Patten **Location: cm1/11E09.** **Art Unit: 1654** **Friday, June 13, 2003** **Case Serial Number: 830964**	**Edward Hart** **Location: Biotech-Chem Library** **CM1-6B02** **Phone: 305-9203** **edward.hart@uspto.gov**

Search Notes

Examiner Patten,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

STIC-Biotech/ChemLib 96531

From: Patten, Patricia
Sent: Friday, June 13, 2003 11:45 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search

Please search SEQ ID No. 1 of 09/830,964

Thank you.

Patricia Patten

1654 CM1 11E09

308-1189

Edward Hart
Technical Info. Specialist
STIC/Biotech
CM1 6B02 Tel: 305-9203

Searcher:_____
Phone:_____
Location:_____
Date Picked Up: 6/13/03
Date Completed: 6/13/03
Searcher Prep/Review:_____
Clerical:_____
Online time:_____

TYPE OF SEARCH:
NA Sequences:_____
AA Sequences:_____
Structures:_____
Bibliographic:_____
Litigation:_____
Full text:_____
Patent Family:_____
Other:_____

VENDOR/COST (where applic.)
STN:_____
DIALOG:_____
Questel/Orbit:_____
DRLink:_____
Lexis/Nexis:_____
Sequence Sys.: 0
WWW/Internet:_____
Other (specify):_____